Volume 650

V(D)J Recombination

Edited by
Pierre Ferrier

V(D)J Recombination

ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY

Editorial Board:
NATHAN BACK, *State University of New York at Buffalo*
IRUN R. COHEN, *The Weizmann Institute of Science*
ABEL LAJTHA, *N.S. Kline Institute for Psychiatric Research*
JOHN D. LAMBRIS, *University of Pennsylvania*
RODOLFO PAOLETTI, *University of Milan*

Recent Volumes in this Series

Volume 642
THE SARCOMERE AND SKELETAL MUSCLE DISEASE
Edited by Nigel G. Laing

Volume 643
TAURINE 7
Edited by Junichi Azuma

Volume 644
TROPOMYOSIN
Edited by Peter Gunning

Volume 645
OXYGEN TRANSPORT TO TISSUE XXX
Edited by Per Liss, Peter Hansell, Duane F. Bruley, and David K. Harrison

Volume 646
EARLY NUTRITION PROGRAMMING AND HEALTH OUTCOMES IN LATER LIFE
Edited by Berthold Koletzko, Tamás Desci, Denes Molnár, and Anne De la Hunty

Volume 647
THERAPEUTIC TARGETS OF THE TNF SUPERFAMILY
Edited by Iqbal Grewal

Volume 648
ARTERIAL AND ALLIED CHEMORECEPTORS
Edited by Constancio Gonzalez, Colin A. Nurse, and Chris Peers

Volume 649
MOLECULAR MECHANISMS OF SPONDYLOARTHROPATHIES
Edited by Carlos López-Larrea, and Roberto Díaz-Peña

Volume 650
V(D)J RECOMBINATION
Edited by Pierre Ferrier

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

V(D)J Recombination

Edited by

Pierre Ferrier, MD, PhD

Centre d'Immunologie de Marseille-Luminy, Université d'Aix Marseille, Marseille, France

Springer Science+Business Media, LLC

Landes Bioscience

Springer Science+Business Media, LLC
Landes Bioscience

Copyright ©2009 Landes Bioscience and Springer Science+Business Media, LLC

All rights reserved.
No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system; for exclusive use by the Purchaser of the work.

Printed in the USA.

Springer Science+Business Media, LLC, 233 Spring Street, New York, New York 10013, USA
http://www.springer.com

Please address all inquiries to the publishers:
Landes Bioscience, 1002 West Avenue, Austin, Texas 78701, USA
Phone: 512/ 637 6050; FAX: 512/ 637 6079
http://www.landesbioscience.com

The chapters in this book are available in the Madame Curie Bioscience Database.
http://www.landesbioscience.com/curie

V(D)J Recomination, edited by Pierre Ferrier. Landes Bioscience / Springer Science+Business Media, LLC dual imprint / Springer series: Advances in Experimental Medicine and Biology.

ISBN: 978-1-4419-0295-5

While the authors, editors and publisher believe that drug selection and dosage and the specifications and usage of equipment and devices, as set forth in this book, are in accord with current recommendations and practice at the time of publication, they make no warranty, expressed or implied, with respect to material described in this book. In view of the ongoing research, equipment development, changes in governmental regulations and the rapid accumulation of information relating to the biomedical sciences, the reader is urged to carefully review and evaluate the information provided herein.

Library of Congress Cataloging-in-Publication Data

V(D)J recombination / edited by Pierre Ferrier.
p. ; cm. -- (Advances in experimental medicine and biology ; v. 650)
Includes bibliographical references and index.
ISBN 978-1-4419-0295-5
1. Genetic recombination. 2. DNA-binding proteins. I. Ferrier, Pierre, 1951- II. Series: Advances in experimental medicine and biology ; v. 650.
[DNLM: 1. Recombination, Genetic. 2. VDJ Recombinases. 3. DNA-Binding Proteins. 4. Recombinant Proteins. W1 AD559 v.650 2009 / QU 475 V393 2009]
QH443.V35 2009
571.9'648--dc22

2009011105

PREFACE

V(D)J recombination: for the community of immunologists and developmental biologists, the molecular route by which B and T lymphocytes acquire their unique function of affording adaptive immunity. Yet, for many—from experienced scientists to trainees—it represents a (rather too) sophisticated process whose true insight is excessively demanding. However, when not simply considered as a private ground for a few aficionados, it can be seen as a way of understanding how mature lymphocytes carry on their basic functions. For the group of aficionados—which includes this editor—it is an elegant paradigm featuring many fascinating evolutionary achievements of which the biological world alone has the secret. These include a subtle biochemical principle most likely hijacked some 470 million years ago from an ancestral gene invader and since then cleverly adapted by jawed vertebrates to precisely cleave and rearrange their antigen receptor (Ig and TCR) loci. This invader would itself have assigned the services of the nonhomologous end joining (NHEJ) DNA repair machinery as well as various DNA polymerases or transferases to work in concert with developmental clues in lymphoid cell lineages to generate an immune repertoire and efficient host surveillance while avoiding autoimmunity.

Recently, important new refinements in these systems have emerged, continuing to challenge our knowledge and beliefs. These are just the topics covered by the senior authors—all established leaders in this field—and their colleagues, whilst writing the various chapters in *V(D)J Recombination.* They lead us through the latest findings concerning the biochemical properties of the V(D)J recombinase (Swanson), its buried and potentially harmful transposase and translocase activities (Oettinger; Roth), the increasing importance of NHEJ, whose dysfunction causes severe forms of immune deficiencies (de Villartay), and the numerous facets in the control of gene rearrangement via non-coding RNA transcription and exquisitely regulated changes in chromosomal structure (Corcoran; Feeney; Jouvin-Marche; Krangel; Oltz and Spicuglia).

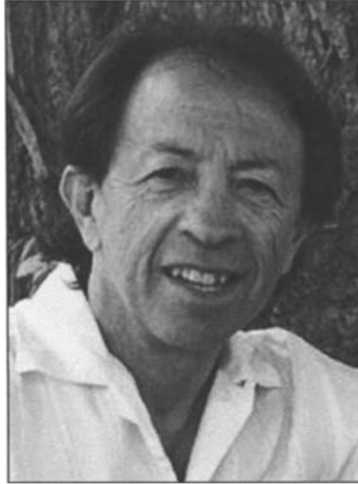
Burning progress on regulatory aspects has included the large-scale dynamics and nuclear compartmentalization of Ig and TCR loci (Singh), the anticipated—but difficult to ascertain—role of dedicated transcription factors (Zhang), the relationships between structural properties of the recombination core apparatus and its cell cycle phase-dependant accumulation/degradation or connection to the chromatin

template (Desiderio), the evolution of these regulatory aspects throughout the phylogeny (Hsu), and how abnormalities in the recombination apparatus/process can contribute to lymphoid malignancies (Macintyre).

Overall, *V(D)J Recombination* represents a tour over this, in all respects, vital process and I would like to greatly acknowledge the efforts of these eminent colleagues for concisely describing its so many aspects. We believe that every advance in this field contributes to strengthening knowledge of fundamental importance both academically and clinically. Together, we hope that the result is an attractive book which will captivate its readers and encourage some to pursue further digging in this seemingly inexhaustible mine of biological resources.

Pierre Ferrier, MD, PhD

ABOUT THE EDITOR...



PIERRE FERRIER is a Principal Investigator and Research Director at the Centre d'Immunologie de Marseille-Luminy (CIML), France. He has also worked as a Director of Marseille-Nice Genopole, a local consortium of more than twenty laboratories aimed at developing high-throughput research techniques in genomics. Main research interests include the analysis of the molecular mechanisms responsible for the control of gene expression and recombination programs during hematopoietic cell development and pathogenesis. He is a member of several national and international scientific organizations including the Institut National de la Santé et de la Recherche Médicale (Inserm), the Agence Nationale de la Recherche (ANR), the Association pour la Recherche sur le Cancer (ARC), the Human Frontier Science Program Organization (HFSPO), and the Université Virtuelle Médicale de Monaco (UVMM). Pierre Ferrier received his academic degrees from Montpellier (MD) and Marseille (PhD) Universities, France. He was a post-doctoral fellow (1986-90) in the laboratory of Prof. F.W. Alt at the Columbia University College of Physicians and Surgeons, New York, NY, USA.

PARTICIPANTS

Iratxe Abarrategui
Centre for Epigenetics
Biotech Research and Innovation Centre
Copenhagen
Denmark

Vahid Asnafi
INSERM EMI U0210
Hopital Necker-Enfants Malades
AP-HP
University Paris-Descartes
Paris
France

Kheïra Beldjord
INSERM EMI U0210
Hopital Necker-Enfants Malades
AP-HP
University Paris-Descartes
Paris
France

Daniel J. Bolland
Laboratory of Chromatin and Gene Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Marie Bonnet
Centre d'Immunologie de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Vicky L. Brandt
Department of Pathology and Program in Molecular Pathogenesis
The Helen L. and Martin S. Kimmel Center for Biology and Medicine
Skirball Institute for Biomolecular Medicine
New York University School of Medicine
New York, New York
USA

Robin Milley Cobb
Department of Microbiology and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Anne E. Corcoran
Laboratory of Chromatin and Gene Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Saïda Dadi
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
Marseille, France
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Stephen Desiderio
Department of Molecular Biology
and Genetics
Institute for Cell Engineering
The Johns Hopkins University
School of Medicine
Baltimore, Maryland
USA

Jean-Pierre de Villartay
INSERM, U768, Unité Développement
Normal
et Pathologique du Système Immunitaire
and
Faculté de Médecine René Descartes
Université Paris-Descartes
and
AP-HP, Hôpital Necker Enfants Malades
Service d'Immunologie et
d'Hématologie Pédiatrique
Paris
France

Ann J. Feeney
The Scripps Research Institute
Department of Immunology
La Jolla, California
USA

Pierre Ferrier
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Patrizia Fuschiotti
Department of Immunology
University of Pittsburgh
School of Medicine
Pittsburgh, Pennsylvania
USA

Ellen Hsu
Department of Physiology
and Pharmacology
State University of New York
Health Science Center at Brooklyn
Brooklyn, New York
USA

Kristen Johnson
Department of Molecular Genetics
and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

Mary Elizabeth Jones
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

Evelyne Jouvin-Marche
INSERM, U823
Faculté de Médecine
Institut d'Oncologie/Developpement
Albert Bonniot et Institut Français
du Sang
Université Joseph Fourier-Grenoble I
Grenoble
France

Michael S. Krangel
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

Sushil Kumar
Department of Medical Microbiology
and Immunology,
Creighton University Medical Center
Omaha, Nebraska
USA

Sandrine Le Noir
INSERM EMI U0210
Hôpital Necker-Enfants Malades
AP-HP
Université Paris-Descartes
Paris
France

Yun Liu
Department of Molecular Biology and Genetics
Institute for Cell Engineering
The Johns Hopkins University School of Medicine
Baltimore, Maryland
USA

Elizabeth A. Macintyre
INSERM EMI U0210
Hôpital Necker-Enfants Malades
AP-HP
Université Paris-Descartes
Paris
France

Patrice Noël Marche
INSERM, U823
Faculté de Médecine
Institut d'Oncologie/Developpement
Albert Bonniot et Institut Français du Sang
Université Joseph Fourier-Grenoble I
Grenoble
France

Adam G.W. Matthews
Department of Molecular Biology
Massachusetts General Hospital
and
Department of Genetics
Harvard Medical School
Boston, Massachusetts
USA

Marjorie A. Oettinger
Department of Molecular Biology
Massachusetts General Hospital
and
Department of Genetics
Harvard Medical School
Boston, Massachusetts
USA

Eugene M. Oltz
Department of Microbiology and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Prafulla Raval
Department of Medical Microbiology and Immunology
Creighton University Medical Center
Omaha, Nebraska
USA

Karen L. Reddy
Department of Molecular Genetics and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

David B. Roth
Department of Pathology and Program in Molecular Pathogenesis
The Helen L. and Martin S. Kimmel Center for Biology and Medicine
Skirball Institute for Biomolecular Medicine
New York University School of Medicine
New York, New York
USA

Harinder Singh
Department of Molecular Genetics and Cell Biology
Howard Hughes Medical Institute
The University of Chicago
Chicago, Illinois
USA

Salvatore Spicuglia
Centre d'Immunologie
de Marseille-Luminy
Université d'Aix Marseille
CNRS, UMR 6102
Inserm, U 631
Marseille
France

Patrick C. Swanson
Department of Medical Microbiology
and Immunology
Creighton University Medical Center
Omaha, Nebraska
USA

Lance R. Thomas
Department of Microbiology
and Immunology
Vanderbilt University
Nashville, Tennessee
USA

Andrew L. Wood
Laboratory of Chromatin and Gene
Expression
Babraham Institute
Babraham Research Campus
Cambridge
UK

Li Zhang
Department of Molecular Biology
and Genetics
Institute for Cell Engineering
The Johns Hopkins University
School of Medicine
Baltimore, Maryland
USA

Yuan Zhuang
Department of Immunology
Duke University Medical Center
Durham, North Carolina
USA

CONTENTS

1. EARLY STEPS OF V(D)J REARRANGEMENT: INSIGHTS FROM BIOCHEMICAL STUDIES OF RAG-RSS COMPLEXES 1

Patrick C. Swanson, Sushil Kumar and Prafulla Raval

Abstract ... 1
Introduction ... 1
Assembly and Organization of Single Site and Synaptic RAG-RSS Complexes ... 3
Insights into RAG-Mediated RSS Recognition and Cleavage Mechanisms ... 5
Elements Guiding Enforcement of the 12/23 Rule ... 8
Transcription Factor-Assisted Targeting of Antigen Receptor Loci ... 10
Conclusion and Future Directions ... 11

2. REGULATION OF RAG TRANSPOSITION ... 16

Adam G.W. Matthews and Marjorie A. Oettinger

Abstract ... 16
Introduction ... 16
Biochemistry of V(D)J Recombination ... 16
Overview of RAG Transposition ... 19
Regulation of RAG Transposition ... 24
Current Understanding of How RAG Transposition Is Regulated ... 24
Additional Potential Regulatory Mechanisms ... 25
Conclusion ... 27

3. RECENT INSIGHTS INTO THE FORMATION OF RAG-INDUCED CHROMOSOMAL TRANSLOCATIONS ... 32

Vicky L. Brandt and David B. Roth

Abstract ... 32
Introduction ... 32
Overview of the V(D)J Recombination Reaction ... 33
Potential Mechanisms of RAG-Mediated Translocations ... 34
Mistaken Identities: Substrate Selection Errors ... 34
The Ends That Got Away: Errors in Joining ... 36

4. V(D)J RECOMBINATION DEFICIENCIES 46

Jean-Pierre de Villartay

Abstract 46
Introduction 46
***RAG1* and *RAG2* Deficiencies** 47
T-B-SCID with Radiosensitivity 50

5. LARGE-SCALE CHROMATIN REMODELING AT THE IMMUNOGLOBULIN HEAVY CHAIN LOCUS: A PARADIGM FOR MULTIGENE REGULATION 59

Daniel J. Bolland, Andrew L. Wood and Anne E. Corcoran

Abstract 59
Introduction 60
Chromatin Remodeling 62
Intergenic Transcription 63
Intergenic Transcription in the Mouse *Igh* Locus V Region 63
Antisense Transcription 64
Antisense Transcription in the *Igh* Locus V Region 64
Antisense and Intergenic Transcription in the *Igh* D Region 66
Subnuclear Relocalisation 66
3-Dimensional Alterations in Chromatin Structure 67
Transcription Factories 68
Biased Recombination Frequency Explainedby Numerous Mechanisms 68
Allelic Choice and Allelic Exclusion 68
Other Antigen Receptor Loci 69
Future Directions 69

6. GENETIC AND EPIGENETIC CONTROL OF V GENE REARRANGEMENT FREQUENCY 73

Ann J. Feeney

Abstract 73
Introduction 73
Sequence Variation in RSS Can Greatly Affect Recombination 74
RSS Is Not Always Responsible for Unequal Rearrangement 75
Chromatin as the Gatekeeper of Accessibility 75
Role of Transcription Factors in Controlling Rearrangement 77
Conclusion 79

7. DYNAMIC ASPECTS OF TCRα GENE RECOMBINATION: QUALITATIVE AND QUANTITATIVE ASSESSMENTS OF THE TCRα CHAIN REPERTOIRE IN MAN AND MOUSE 82

Evelyne Jouvin-Marche, Patrizia Fuschiotti and Patrice Noël Marche

Abstract 82

Introduction **82**
Complexity of Mouse and Human TCRAD Locus **83**
Analysis of Human and Mouse TCRA-Chain Diversity **84**
Comparison between the Frequencies of Rearrangements in Thymus and Peripheral T-Lymphocytes **85**
The Size of the Mouse and Human TCRα Repertoire **87**
Conclusion **90**

8. GERMLINE TRANSCRIPTION: A KEY REGULATOR OF ACCESSIBILITY AND RECOMBINATION **93**

Iratxe Abarrategui and Michael S. Krangel

Abstract **93**
Introduction **93**
A Brief History of Germline Transcription and V(D)J Recombination **94**
Disruption of Chromatin by Transcription **95**
Regulation of V(D)J Recombination by Transcription **97**
Future Directions **99**

9. DYNAMIC REGULATION OF ANTIGEN RECEPTOR GENE ASSEMBLY **103**

Lance R. Thomas, Robin Milley Cobb and Eugene M. Oltz

Abstract **103**
Introduction **103**
Developmental Control of V(D)J Recombination **104**
Genetic Control of Recombinase Accessibility **105**
Chromatin Accessibility Control Mechanisms for V(D)J Recombination **107**
Control of V(D)J Recombination by Nuclear Compartmentalization **109**
Primary Activation of Antigen Receptor Loci for D to J Rearrangement **109**
Long-Range Control of V(D)J Recombination **111**
Allelic Exclusion **111**
Conclusion **113**

10. MOLECULAR GENETICS AT THE T-CELL RECEPTOR β LOCUS: INSIGHTS INTO THE REGULATION OF V(D)J RECOMBINATION **116**

Marie Bonnet, Pierre Ferrier and Salvatore Spicuglia

Abstract **116**
Introduction **116**
Overview of the *Tcrb* Genomic Structure and Recombination Properties **117**
***Tcrb*-RSSs and Rearrangement Efficiency** **117**
***Cis*-Regulatory Elements at the *Tcrb* Locus** **119**
Trans-Regulators of *Tcrb* Locus Expression/Recombination **122**
Chromatin Accessibility **123**
Allelic Exclusion at the *Tcrb* Locus **126**
Conclusion and Future Direction **128**

11. MOLECULAR PATHWAYS AND MECHANISMS REGULATING THE RECOMBINATION OF IMMUNOGLOBULIN GENES DURING B-LYMPHOCYTE DEVELOPMENT 133

Kristen Johnson, Karen L. Reddy and Harinder Singh

Abstract 133
Introduction 133
B-Cell Fate Specification and the Joining of D-to-J_H Segments 137
B-Cell Fate Commitment and V-to-DJ_H Rearrangement 138
The Pre-B-Cell Checkpoint and the Induction of Light-Chain Recombination 141
Allelic Exclusion 143
Perspectives 144

12. REGULATION OF V(D)J RECOMBINATIONBY E-PROTEIN TRANSCRIPTION FACTORS 148

Mary Elizabeth Jones and Yuan Zhuang

Abstract 148
Introduction 148
Transcriptional Control of Ig and TCR Antigen Receptor and Their Associated Genes 149
Induction of Ig and TCR Gene Rearrangement 150
Regulation of the Developmental Window for V(D)J Recombination 150
Conclusion 153

13. TEMPORAL AND SPATIAL REGULATION OF V(D)J RECOMBINATION: INTERACTIONS OF EXTRINSIC FACTORS WITH THE RAG COMPLEX 157

Yun Liu, Li Zhang and Stephen Desiderio

Abstract 157
Functional Organization of RAG-1 and RAG-2 157
Temporal Regulation of V(D)J Recombination through Interactions with the RAG-2 Non-Core Region 158
Locus Specificity: General Remarks 159
Epigenetic Modifications of Possible Relevance to V(D)J Recombination 159
DNA Methylation 159
Nucleosome Phasing 159
Histone Acetylation 160
Histone H3 K9 Methylation 160
Histone H3 K4 Methylation 160
Direct Recognition of Modified Histone H3 by the V(D)J Recombinase 160
Evidence for Allosteric Regulation of V(D)J Recombinase Activity by Histone H3 Trimethylated at Lysine 4 161
Future Directions: Deposition and Integration of Epigenetic Signals Controlling V(D)J Recombination 162

14. V(D)J RECOMBINATION: OF MICE AND SHARKS 166

Ellen Hsu

Abstract 166
Introduction 166
V(D)J Rearrangement 168
V(D)J Rearrangement Patterns 172
Rabbit 172
Multiple IgH Loci in Other Vertebrate Species 173
Conclusion 176

15. NORMAL AND PATHOLOGICAL V(D)J RECOMBINATION: CONTRIBUTION TO THE UNDERSTANDING OF HUMAN LYMPHOID MALIGNANCIES 180

Saïda Dadi, Sandrine Le Noir, Vahid Asnafi, Kheïra Beldjord and Elizabeth A. Macintyre

Abstract 180
Introduction 180
Diagnostic Clonality Analysis 181
Recombinase Mediated Oncogenesis 185
Conclusion 190

INDEX 195

Chapter 1

Early Steps of V(D)J Rearrangement: Insights from Biochemical Studies of RAG-RSS Complexes

Patrick C. Swanson,* Sushil Kumar and Prafulla Raval

Abstract

V(D)J recombination is initiated by the synapsis and cleavage of a complementary (12/23) pair of recombination signal sequences (RSSs) by the RAG1 and RAG2 proteins. Our understanding of these processes has been greatly aided by the development of in vitro biochemical assays of RAG binding and cleavage activity. Accumulating evidence suggests that synaptic complex assembly occurs in a step-wise manner and that the RAG proteins catalyze RSS cleavage by mechanisms similar to those used by bacterial transposases. In this chapter we will review the molecular mechanisms of RAG synaptic complex assembly and 12/23-regulated RSS cleavage, focusing on recent advances that shed new light on these processes.

Introduction

The antigen-binding variable domains of immunoglobulins and T-cell receptors exhibit great structural diversity that mostly originates from a site-specific DNA rearrangement process, called V(D)J recombination, that assembles the exons encoding the variable domains of these proteins from germline variable (V), diversity (D) and joining (J) gene segments during lymphocyte development.[1] Adjacent to each gene segment lies a recombination signal sequence (RSS); each RSS contains a conserved heptamer and nonamer motif (consensus heptamer: 5'-CACAGTG-3'; consensus nonamer: 5'-ACAAAAACC-3') separated by "spacer" DNA, normally 12 base pairs (bp) or 23 bp long (12-RSS and 23-RSS, respectively), which displays some sequence preferences proximal to the heptamer[2] but is otherwise not well conserved. V(D)J recombination is generally directed between two gene segments with different RSSs, a restriction termed the 12/23 rule that serves to facilitate productive receptor gene assembly.

The biochemistry of V(D)J recombination can be conceptually divided into a cleavage phase and a joining phase (Fig. 1). To initiate the cleavage phase, two lymphoid cell-specific proteins encoded by recombination activating gene-1 and -2 (RAG1 and RAG2, respectively[3,4]), possibly assisted by high mobility group proteins of the HMG-box family (HMGB1 and HMGB2, called HMGB1/2 henceforth; discussed further below), bring two different gene segments into close proximity through interactions with the adjoining 12- and 23-RSS (forming a "synaptic" complex) and then catalyze a DNA double-strand break (DSB) at each RSS between the heptamer and the coding segment.[5,6] RAG-mediated cleavage produces two types of DNA ends: blunt and 5'-phosphorylated signal ends containing the RSS and coding ends covalently sealed as DNA hairpins.[7,8] These reaction intermediates originate from a two-step cleavage mechanism in which

*Corresponding Author: Patrick C. Swanson—Department of Medical Microbiology and Immunology, Creighton University Medical Center, Omaha, Nebraska 68178, USA. Email: pswanson@creighton.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

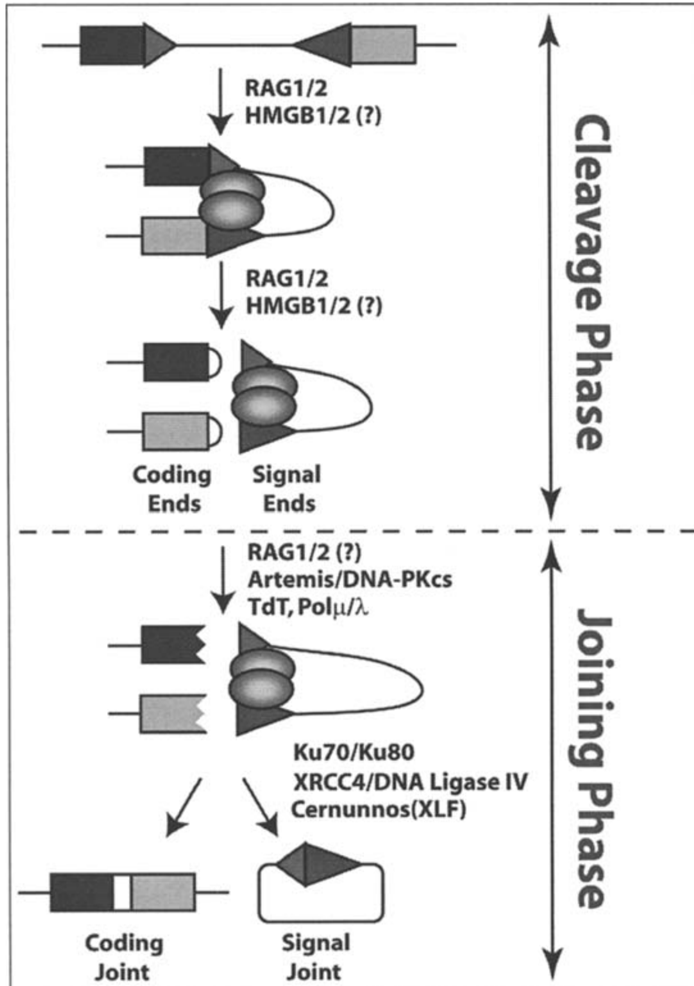


Figure 1. Overview of V(D)J recombination (adapted from Fugmann et al[6]). In the cleavage phase of V(D)J recombination, coding segments (filled rectangles), flanked by a 12-RSS or 23-RSS (small or large triangles, respectively) are assembled into a synaptic complex by the RAG proteins, possibly assisted by HMGB1/2 (filled ovals). Coupled cleavage by the RAG proteins yields blunt signal ends and coding ends sealed as DNA hairpins. In the joining phase of V(D)J rearrangement, sealed coding ends are resolved by an Artemis/DNA-PKcs complex and may be further processed by TdT (if present) and DNA polymerases μ and/or λ (Polμ/λ). Processed coding ends are joined to create imprecise coding joints that may have gained palindromic (P) or nontemplated (N) nucleotides through asymmetric hairpin opening or TdT-mediated addition, respectively, or lost nucleotides through end processing reactions (open rectangle). Signal ends are joined to create signal joints that are typically precise. Alternative, less frequent joining events, such as open-shut and hybrid joints are not shown for simplicity. Signal and coding joint formation is mediated by the NHEJ pathway, which includes Ku70, Ku80, XRCC4, DNA Ligase IV and Cernunnos (XLF). Although the processing and joining reactions are shown as sequential processes, these steps may be integrated and iterative for joining of incompatible coding ends, involving single-strand ligation, processing of the unligated strand by Artemis/DNA-PKcs and DNA polymerases and eventual ligation of the second strand resulting in repaired double-stranded DNA.[101]

the RAG proteins first nick the RSS at the 5' end of the heptamer and then use the resulting 3'-OH to catalyze a direct transesterification reaction on the opposing phosphodiester bond.[9] In the joining phase, the two signal ends are typically ligated precisely, forming a signal joint, and the coding ends are subjected to reactions that resolve the hairpins and then process and connect the DNA ends to form coding joints. As a result, coding joints often show evidence of nucleotide gain or loss at the coding ends. Infrequently, alternative outcomes of V(D)J recombination are observed in which one gene segment is joined to the RSS of another gene segment ("hybrid joint") or is separated and rejoined to the same RSS ("open-shut joint").[10,11] Efficient signal and coding joint formation requires a competent nonhomologous end-joining (NHEJ) repair pathway, including Ku70, Ku80, XRCC4, DNA Ligase IV and XLF/Cernunnos.[12-14] Coding joint formation requires two additional factors not strictly essential for joining signal ends, Artemis and DNA-PKcs, which together function as a structure-specific endonuclease responsible for opening the DNA hairpins on coding ends.[15] Asymmetric hairpin opening can give rise to palindromic (P) nucleotides being inserted in coding joints. Terminal deoxynucleotidyl transferase (TdT) and DNA polymerases μ and/or λ (Pol μ/λ) can further diversify these junctional sequences by catalyzing addition of nontemplated (N) nucleotides to coding ends (TdT) and processing incompatible DNA ends to facilitate end-joining (Pol μ/λ).[16,17] A detailed consideration of the proteins involved in the processing and repair of V(D)J recombination intermediates is beyond the scope of this review, but has been discussed elsewhere.[18-20]

Here we review and discuss the molecular mechanisms of V(D)J recombination, focusing on the cleavage phase of this process and emphasizing new insights. Readers are referred to previous reviews for more detailed discussion of early studies of RAG protein biochemistry, including the establishment of cell-free assays of V(D)J cleavage and joining[5,6] and the identification and characterization of the various structural domains of the RAG proteins.[21]

Assembly and Organization of Single Site and Synaptic RAG-RSS Complexes

Cell-free assays of V(D)J cleavage established using truncated, catalytically active "core" forms of RAG1 (full-length 1040 a.a.; core residues 384-1008) and RAG2 (full-length 517 a.a.; core residues 1-387) demonstrated that the RAG1/2 complex is both necessary and sufficient to mediate RSS cleavage[9] and that RAG cleavage activity exhibits metal ion-dependence: Mn^{2+} supports RAG-mediated cleavage of a single RSS, whereas Mg^{2+} is required for coupled cleavage of RSS pairs abiding by the 12/23 rule.[22,23] In natural progression, later studies identified and characterized discrete RAG-RSS complexes with increasing complexity, with early work focused on RAG complexes assembled on a single RSS and later work analyzing higher-order RAG synaptic complexes. Most of this work has been reviewed and discussed elsewhere.[5,6,24] Therefore, only salient features will be highlighted here.

Core RAG1 contains three structurally distinct regions:[21] an amino-terminal nonamer binding domain (NBD, residues 389-442) that interacts with the RSS nonamer,[25,26] a central domain (residues 528-760) that recognizes the heptamer and exhibits single-strand DNA binding activity and a C-terminal domain (residues 761-979) that binds double-stranded DNA nonspecifically and cooperatively. Core RAG1 alone exists in solution primarily as a stable dimer[27-29] and binds an isolated RSS with moderate affinity (Kd ~41 nM)[28] as a dimer[27,28,30] (although higher-order aggregates are detectable at elevated RAG1 concentrations and conditions of low ionic strength[31]) whereas RAG2 is predominantly monomeric in solution[29] and shows little, if any DNA binding activity.[25,26,32-34] RAG1 and RAG2 interact with one another in the absence of DNA[27,29,35] and together bind a single RSS with greater specificity than RAG1 alone.[32,33,36] Purified core RAG1/2 proteins variably assemble one[29,32,33] or two[34,37] major protein-DNA complexes detectable using an electrophoretic mobility shift assay (EMSA). The relative abundance of these complexes, now generally called SC1 and SC2 (for "single RSS complex"), depends partly on how the RAG proteins are expressed and purified:[37,38] in our laboratory, individually expressed and purified RAG proteins tend to assemble only SC1, coexpressed RAG proteins purified under high salt conditions form

more SC1 than SC2 and coexpressed RAG proteins purified using milder conditions predominantly assemble SC2. Both complexes possess similar intrinsic cleavage activity,[34,37] but differ in RAG protein stoichiometry. Swanson reported that both complexes contain a RAG1 dimer, but incorporate either one (SC1) or two (SC2) RAG2 molecules.[37] Mundy et al reported comparable results for RAG2 in these complexes, but presented evidence suggesting SC1 and SC2 contain three or more RAG1 subunits.[34] Possible explanations for this apparent discrepancy have been discussed previously[24] and will not be revisited here, but we note that recent data reported by De et al provides corroborating evidence supporting the contention that RAG1 exists as a dimer in an SC (RAG2 stoichiometry was not determined).[39] The tetrameric RAG1/RAG2 configuration reported for SC2 is also consistent with data published by Bailin et al.[29]

Mutagenesis studies[40-42] revealed that RAG1 contains three carboxylate residues (asp-600, asp-708 and glu-962) critical for catalysis that resemble a "DDE motif" found in many transposases and integrases.[43] Similar to the Tn5 transposase,[44,45] biochemical studies established that a single RAG1 subunit contributes all three carboxlate residues to single active site which mediates sequential nicking and hairpin formation steps of the cleavage reaction[46,47] and that these reactions are catalyzed in trans; that is, by the subunit of the RAG1 heterodimer not bound to the nonamer of the RSS being cleaved.[47]

While the RAG proteins themselves are sufficient for assembling SC1 and SC2, HMGB1/2 proteins are known to facilitate RAG-mediated binding and cleavage of an isolated 23-RSS, but not a 12-RSS, in vitro.[48] The RAG proteins also require the presence of HMGB1/2 to efficiently assemble a complex containing a complementary (12/23) pair of RSSs ("paired complex" or PC) and mediate coupled cleavage at both RSSs adhering to the 12/23 rule in vitro.[48,49] Whether HMGB1/2 also assist the RAG proteins during V(D)J recombination in vivo has not been formally established nor entirely ruled out,[50] since HMGB1/2 exhibit functional redundancy in RAG binding and cleavage assays.[51] The HMGB1/2 proteins are nonhistone chromosomal DNA binding proteins known to promote DNA bending and facilitate assembly of nucleoprotein complexes;[52] HMGB1 further functions as an alarmin to signal cellular damage in response to inflammatory processes.[53] HMGB1/2 proteins contain tandem homologous HMG-box domains (called A and B) attached to a basic linker and an acidic tail. HMGB1/2 interacts with the NBD of RAG1 in the absence of DNA and enhances the intrinsic DNA bending activity of the RAG proteins.[54] The integration of HMGB1/2 into RAG-RSS complexes can often be detected as a supershift by EMSA.[51,55] Recent structure-function studies conducted in our laboratory[56,57] suggest that both HMG-box domains must be competent to bend DNA and physically linked together in either orientation (AB or BA) to stimulate RAG-mediated 23-RSS cleavage in the presence of Mg^{2+}. Interestingly, single HMG-box domains can be integrated into 23-RSS-RAG complexes,[56-58] but cannot stimulate 23-RSS cleavage unless Mn^{2+} replaces Mg^{2+} in the reaction,[57,58] or 12-RSS partner is added to promote synapsis.[57] These results suggest the two HMG-box domains have separable but potentially redundant roles in stimulating RAG binding and cleavage activity in vitro and that synapsis promotes a conformational change that bypasses the need for one of these domains. HMGB1 lacking the acidic tail stimulates RAG binding and cleavage activity at lower concentrations than full-length HMGB1, but promotes aggregation of RAG-RSS complexes.[56,58] Moreover, loss of the acidic tail enables HMGB1 mutants that otherwise fail to support RAG-mediated synapsis to stimulate PC formation.[56] These data suggest the acidic tail helps maintain the correct oligomerization state of RAG synaptic complexes. The acidic tail is also known to facilitate HMGB1-mediated nucleosome repositioning,[59,60] which may help promote RSS accessibility in nucleosomal DNA.[61-63]

Synaptic complex assembly is thought to proceed via initial formation of SC2 followed by capture of an appropriate partner RSS to form a PC. This "capture model" of assembly was suggested initially by biochemical experiments demonstrating that SC2 can be driven to form the PC by adding appropriate partner RSS[34] and the observation that RAG cleavage activity is greater when synaptic complexes are assembled in step-wise fashion by adding free 23-RSS to a 12-RSS-RAG complex (or vice versa) than when they are assembled by mixing preformed 12-RSS-RAG and

23-RSS-RAG complexes together.[64] This model has gained in vivo experimental support from a recent study by Curry et al[65] showing that nicks can be detected at endogenous 12-RSSs, but not at 23-RSSs, in lymphoid cells. These findings lead the authors to propose a model in which RAG proteins bind and nick a 12-RSS first, then capture and nick a 23-RSS and, in rapid succession, finally cleave both RSSs. This model is consistent with previous biochemical studies showing that nicking can occur on an RSS in the absence of synapsis,[66,67] but nicking at one RSS is required for efficient cleavage of its partner.[22,66] The capture model is also consistent with data this laboratory and others have published showing that the complement of RAG proteins is the same between a RAG complex bound to a single RSS (as SC2) and the PC.[34,37] Interestingly, these studies show that molecules of RAG2, but not RAG1, freely re-assort during PC assembly.[34,37] Work from this laboratory suggests that the PC contains two molecules each of RAG1 and RAG2 and that this heterotetramer configuration remains the same through the cleavage steps of V(D)J recombination.[37] Another study reported the same stoichiometry for RAG2 in the PC,[34] but others conclude the PC contains three or more RAG1 subunits.[34,46] Possible scenarios to explain these discordant results have been discussed elsewhere.[24]

How are the RSSs arranged in the synaptic complex? Early observations that the efficiency of in vitro coupled cleavage[22] and in vivo V(D)J rearrangement[68] is more sensitive to shortening of the intersignal distance when the RSSs are positioned in an inversional configuration than when they are positioned in a deletional configuration argued that the RSSs are aligned in a parallel, rather than anti-parallel orientation in the synaptic complex. To test this possibility more directly, Cibutaru et al recently measured levels of fluorescence resonance energy transfer (FRET) in RAG synaptic complexes assembled under various conditions on 12- and 23-RSS oligonucleotide substrates labeled with FAM and TAMRA in different configurations.[69] Significant FRET was detected only when the following three conditions were met: (i) the fluorophores were placed on different RSSs (but not the same RSS); (ii) the two RSSs contained different length spacers (i.e., abiding by the 12/23 rule); and, (iii) synaptic complexes were assembled in binding reactions containing Mg^{2+} and the full complement of RAG1/2 and HMGB1/2 proteins. Interestingly, FRET was observed in synaptic complexes regardless of which end of a given RSS was labeled; the only apparent requirement was that the two fluorophores were placed on different RSSs (12 and 23). These data suggest that the distance between the ends of the two bound RSSs in the synaptic complex are approximately the same. Given this constraint and limitations on the maximal distance between fluorophores to observe FRET, the authors propose the two RSSs likely adopt a bent and crossed configuration in the PC.[69]

Insights into RAG-Mediated RSS Recognition and Cleavage Mechanisms

Interactions between the RAG proteins and DNA have been investigated using a variety of approaches and the insights from these studies have greatly improved our understanding of how the RAG proteins recognize and cleave their RSS targets. Much of the early work has been extensively reviewed,[5,6,24] so it will not be covered in depth here. Chemical and DNase I protection and modification interference footprinting assays performed on RAG complexes assembled on a single RSS suggest RAG1 primarily interacts with the nonamer and adjacent spacer sequence, whereas RSS contacts in complexes containing both RAG proteins are overlapping, but more expansive, extending from the nonamer, through the spacer and into the 3' end of the heptamer, with a bias of phosphate contacts toward one face of the DNA helix.[32,70,71] Photo cross-linking studies suggest RAG1 mediates most of the contact with the RSS, with RAG2-RSS interactions more localized to the junction of the heptamer and coding segment.[27,36,72,73] Integration of HMGB1/2 into 23-RSS-RAG complexes enables detection of heptamer-spacer contacts resembling those observed in 12-RSS-RAG complexes that are not otherwise visualized in 23-RSS complexes containing RAG1/2 alone,[51,55] suggesting HMGB1 stabilizes RAG association with the heptamer in these complexes. Ethylation interference footprinting suggests HMGB1/2 contacts the 23-RSS proximal to the nonamer, expanding the footprint of the RAG proteins in this region.[51] Although RAG contacts at the junction of the heptamer and coding sequence are not readily detected in RAG

complexes assembled on a single RSS, this region is protected from DNase I cleavage in synaptic complexes.[74] Nagawa et al showed that synaptic complexes assembled with nicked RSS substrates show slight expansion of the DNase I footprint relative to precleavage synaptic complexes (from ~12 nt to ~16 nt), suggesting that RAG-mediated nicking causes more intimate and stable RAG association with the coding sequence.[75] Pull-down assays showing that nicked RSS substrates are more readily incorporated into synaptic complexes than intact substrates support this contention. Interestingly, two different joining-deficient RAG1 mutants (S723C[76] and K118/9A[77]) were shown to exhibit poor protection of the heptamer-coding junction, leading to speculation that the joining defect is caused by poor coding end retention in the postcleavage synaptic complex.[75] However, close inspection of the mutant RAG1 footprinting patterns in precleavage complexes also reveals that these mutants exhibit less protection of spacer and nonamer sequences compared to wild-type RAG1. This observation argues that these mutations cause a global defect in RAG-RSS complex stability, but can also be interpreted to suggest that the RAG proteins require stable contact with the coding sequence in order to maintain strong interactions with the RSS (or vice versa) in precleavage complexes.

Direct and interference footprinting experiments suggest RAG-RSS complex formation is accompanied by structural distortions in the spacer region and near the site of DNA cleavage.[32,51,70,71] Studies showing that the RAG proteins mediate RSS bending, which is augmented by HMGB1/2,[54] plausibly explain spacer hypersensivity to chemical and enzymatic probes in RAG-RSS complexes. Structural distortions near the cleavage site are likely attributed to base unpairing mediated by the RAG proteins to promote hairpin formation, which is suggested by observations that RAG-mediated RSS cleavage is facilitated by incorporating base-pair mismatches[78,79] or abasic sites[80] at the coding flank. Clues to how these structural distortions may be induced and stabilized are suggested by structural studies of the related Tn5 transposase, which, like the V(D)J recombinase, catalyzes DNA hairpin formation (except that hairpins are formed at the transposon end, which is equivalent to the signal end in V(D)J recombination).[81] Analysis of a Tn5 postcleavage synaptic complex reveals that the transposase promotes extrusion of a thymine from the DNA helix, stabilizing the "flipped base" via stacking interactions with an aromatic tryptophan residue (trp-298).[44] Recent studies indicate a similar mechanism is operative in V(D)J recombination. Two lines of evidence suggest the terminal nucleotide on the bottom strand of the coding flank (C1b, see Fig. 2 inset) is stabilized in an extrahelical configuration by the RAG proteins. First, when thymine is incorporated into the RSS at position C1b, this base exhibits hypersensitivity to permanganate modification under conditions favoring RAG-RSS synaptic complex formation.[82] Second, base removal at C1b potentiates hairpin formation.[80] Both outcomes are consistent with comparable studies of the flipped T2 thymine in the Tn5 transposon end.[83,84] One notable contrast between the two recombination systems is that although the base subjected to flipping in the RSS coding flank and the Tn5 transposon end are both located opposite the nicking site within the hairpin-forming sequence, they are offset from one another by one nucleotide: in the RSS, the base is at the terminus of the sequence; in the Tn5 transposon end, it occupies the penultimate position.

When does base-flipping occur during RSS cleavage? Base-flipping appears to occur after nicking, rather than upon RAG binding to the RSS, as permanganate hypersensitivity is not observed in RAG synaptic complexes assembled on intact substrates.[82] Interestingly, permanganate interference assays reveal that intact substrates bearing oxidized thymine at C1b and S2b are selectively bound by the RAG complex relative to unmodified substrates, with the latter modification being much preferred over the former.[32,51] If the RAG proteins stabilize base-flipping at C1b during the hairpin-forming step, why is prior modification of S2b selected over C1b in interference assays? Since base-flipping is most evident in synaptic complexes assembled on nicked substrates,[82] one possibility is that a conformational change in the RAG complex occurs after synapsis or nicking that alters the position of thymine binding pocket relative to the cleavage site. Thus, an oxidized extrahelical thymine at S2b may be preferentially accommodated over C1b in the binding pocket of a RAG complex bound to an intact RSS. Alternatively, modified S2b may be selected because C1b is more easily flipped if the oxidized base at S2b is already displaced from the DNA helix.

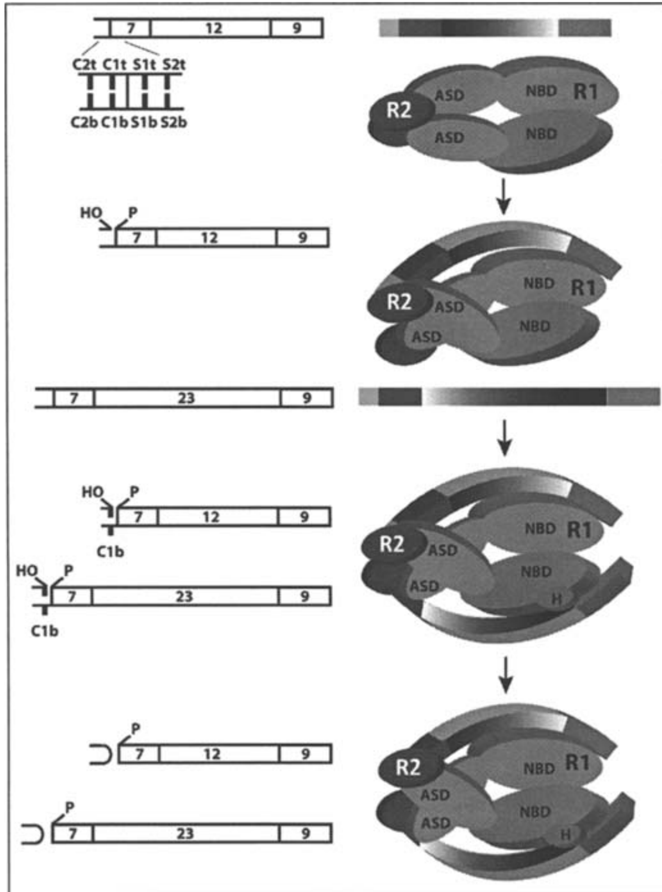


Figure 2. Integrated model of synaptic complex assembly and coupled RSS cleavage. In this simplified scheme, RAG1 (R1) contains an active site domain (ASD) that includes the DDE motif (found within central and C-terminal domains that are not shown[21]) and a nonamer binding domain (NBD). RAG2 (R2) is depicted as a small oval. RAG-RSS complexes are shown at right and reactions catalyzed on the RSSs are diagrammed at left. Nucleotide positions on the top (t) and bottom (b) of the coding (C) and signal (S) sequence at the heptamer-coding junction are also indicated (inset, upper left). RAG1 and RAG2 form a complex, shown here as a heterotetramer based on our work[37] and others[29] (but see text), that preferentially binds a 12-RSS. The RAG complex bends and nicks the 12-RSS at the 5′ end of the heptamer and then captures a 23-RSS to form a PC in which both RSSs are bent and cross over one another. HMGB1/2 (H) may assist in this process at the 23-RSS. The RSSs are shown here wrapping around the outside of the RAG1/2 complex (adapted from Ciubotaru et al[69]). An alternative model in which the RSSs cross over each other on the same face of the protein complex is not shown for simplicity, but is an arrangement that meets constraints imposed by FRET data.[69] Note that the bending and crossing angles shown here are not meant to represent angles derived from experimental measurements. The 23-RSS is nicked in and, in rapid succession, the RAG proteins catalyze hairpin formation at both RSSs by a mechanism involving base-flipping at C1b. RAG-mediated cleavage is shown here catalyzed by a single ASD in trans (i.e., the RAG1 subunit bound to the 12-RSS nonamer cleaves the 23-RSS and vice versa) based on studies of SC1,[47] but this configuration has yet to be formally established for the PC. After cleavage, coding ends are likely released first, with the RAG proteins remaining bound to the signal ends until the signal ends complex is disassembled.

To test whether aromatic residues in RAG1 participate in base-stacking interactions to promote hairpin formation by analogy to Tn5 transposition, two different laboratories performed site-directed mutagenesis of aromatic residues in RAG1.[80,85] Lu et al screened all evolutionarily conserved aromatic residues in the catalytic core of RAG1, selecting mutants failing to support V(D)J cleavage in cells and exhibiting selective impairment of hairpin formation in vitro.[85] The authors identified trp-893 of RAG1 as a plausible candidate for mediating base-stacking interactions, based on the inability of a W893A RAG1 mutant to support hairpin formation and the rescue of this defect by replacing alanine with tyrosine at residue 893 or by introducing mismatched base pairs near the RSS cleavage site. In contrast, a later, more limited mutagenesis study concluded that trp-893 is unlikely to mediate base-stacking because the cleavage defect observed with the W893A RAG1 mutant was found to depend on the coding flank composition.[80] Specifically, Grundy et al showed that RSS substrates containing "bad" coding flanks (5'-GATTC-3' or 5'-TCGAC-3') are cleaved less efficiently by W893A RAG1 than by wild-type RAG1, but wild-type and W893A RAG1 exhibit similar activity using substrates containing "good" coding flanks (5'-ACCTG-3'). Thus, the authors speculated that a trp-893 mutation affects a step following cleavage. However, because the W893A RAG1 mutant supports moderate cleavage of oligonucleotide substrates under conditions favoring synapsis in trans but poor nicking and hairpin formation when the same RSSs are embedded in cis in a PCR-generated substrate,[85] it is also possible that trp-893 mediates protein-protein or protein-DNA interactions to facilitate synaptosome assembly and activity on longer, more physiological substrates that are largely dispensable in reactions performed on oligonucleotide substrates.

Rather than trp-893, Grundy et al argue that trp-956 is a more plausible candidate for stabilizing base-flipping because although a W956A mutant exhibits defects in both nicking and hairpin formation in Mg^{2+} (also reported by Lu et al[85]), W956A RAG1 cleavage activity is substantially rescued by incorporating an abasic site at C1b of the RSS substrate.[80] That the W956A RAG1 mutant is substantially impaired in catalyzing both steps of the cleavage reaction in Mg^{2+} is not the outcome expected based on the precedent set by analysis of its presumed counterpart, W298A Tn5, which exhibits defects in hairpin formation, but not nicking.[83] However, given the close proximity of trp-956 to glu-962, which is required for catalysis,[40,41] a W956A mutation may cause structural alterations in the active site that prevent the RAGs from nicking RSS substrates efficiently. Alternatively, the observation that introducing abasic sites at C1t and C2t of the coding flank (see Fig. 2, inset) blocks the nicking step raises the possibility that trp-956 is involved in both cleavage steps of V(D)J recombination, first to help identify where the nick should be introduced and second, perhaps following a conformational change, to help stabilize the extrahelical base at C1b in preparation for hairpin formation.

Elements Guiding Enforcement of the 12/23 Rule

How the 12/23 rule is enforced at the molecular level still remains somewhat mysterious. As discussed previously,[6] the 12/23 rule is likely enforced both at the level of synapsis and at the point when nicks at both RSSs are converted to DNA double-strand breaks. At the level of synapsis, Jones and Gellert demonstrated that once the RAG proteins bind a 12-RSS in the presence of HMGB1, the complex becomes structurally biased against capturing another 12-RSS and instead exhibits a strong preference for capturing and integrating a 23-RSS into a PC.[64] However, the opposite is not true: RAG proteins bound to a 23-RSS exhibit only a 5-6 fold preference for incorporating a 12-RSS partner over a 23-RSS partner into a PC. The authors speculate that due to the length of the 23-RSS spacer, the RAG proteins bound to this substrate may undergo rapid isomerization between "12-RSS-like" and "23-RSS-like" RAG complexes, enabling the second site to be occupied by either type of RSS, with only modest selectivity for a 12-RSS. Although the authors envisioned bending of the 23-RSS spacer as the means to achieve isomerization,[64] data showing that the RAG proteins can aberrantly nick a 23-RSS in the spacer region at a position equivalent to the 5'-end of the heptamer in a 12-RSS[55,86] raises the possibility that isomerization is alternatively achieved through "catch and release" of 23-RSS heptamer and spacer sequences. The "conformational

locking" model proposed by Jones and Gellert was developed based on experiments using intact RSS substrates, but is equally plausible for a scenario in which the RAG proteins bind and nick a 12-RSS before synapsis, which, as discussed above, is suggested to occur in vivo.[65]

Once bound to a nicked 12-RSS, the RAG proteins must identify an accessible 23-RSS partner in a background of available 12-RSSs (intact or nicked) and other randomly nicked DNA. The conformational locking model provides a framework to discriminate against binding a second 12-RSS, but not a mechanism to do so. One possibility is that the NBD in the RAG1 heterodimer not bound to the 12-RSS may sample incoming DNA sequences for nonamer-like elements. Should it find a suitable sequence, it may bind (modestly) to this motif, enabling sequences at the appropriate distance to be interrogated for the presence of a suitable heptamer. Thus, should a 12-RSS-RAG complex (as SC2) encounter another 12-RSS, the unoccupied RAG1 subunit could bind it via NBD-nonamer interactions, but the heptamer's proximity would not allow it to be specifically engaged by the active site of the RAG complex, causing the RSS to eventually dissociate. Alternatively, if the same 12-RSS-RAG complex encountered a randomly nicked sequence, the active site may bind the nicked DNA weakly, but if the sequence lacks a suitable nonamer-like motif, the DNA would not be fully anchored to the RAG complex via the NBD and therefore would not trigger transesterification. Thus, only when heptamer and nonamer elements are both present and appropriately spaced in the partner RSS would nicking of the partner and subsequent hairpin formation at both RSSs be initiated. What is the critical checkpoint in this process? Nishihara et al showed that base-flipping at C1b is only observed at a nicked 12-RSS when its appropriate partner is bound by the RAG complex.[82] Hence, the decision to base-flip is likely a critical checkpoint in triggering coupled cleavage, as this step provides the conformational changes required to promote transesterification.

What then influences the decision to initiate base-flipping? This decision is likely influenced by how the RAG proteins detect synapsis, as evidenced by the recent identification of gain-of-function RAG1 mutants that exhibit enhanced in vitro RSS cleavage in Mg^{2+} in the absence of synapsis.[82,87] We identified an E649A RAG1 mutant that, relative to wild-type RAG1, exhibits enhanced RAG-mediated hairpin formation in vitro, but does not display increased recombination activity of plasmid V(D)J recombination substrates containing a 12/23 pair of signal sequences in cell culture. However, this mutant does support greater cleavage and recombination of substrates containing a mispaired or unpaired RSS, suggestive of a selective defect in sensing 12/23-regulated synapsis. Whether the E649A RAG1 mutant supports base-flipping in the absence of synapsis has not been tested, but a RAG1 mutant (called HA3) with a similar phenotype was recently found to mediate synapsis-independent base-flipping.[82] It is notable that in both reports, the mutations conferring the gain-of-function phenotype are located proximal to residues of the DDE motif, which suggests that the domain responsible for catalyzing the steps of V(D)J cleavage also plays a key role in sensing 12/23-regulated synapsis and triggering base-flipping at the cleavage site.

Taken together, the data summarized here support a model of RAG synaptic complex assembly and 12/23-regulated cleavage shown in Figure 2 that involves initial binding, bending and nicking of a 12-RSS by the RAG complex, followed by the selective capture and integration of a free 23-RSS into a synaptic complex in which the two RSSs adopt a bent and crossed configuration and finally completed by 23-RSS nicking and facile conversion of nicks at both RSSs into DNA hairpins by a mechanism that involves base-flipping at C1b. The conformational changes required to mediate this process on physiological substrates may be facilitated in part by mechanisms that underwind DNA, as such substrates are cleaved more efficiently by the RAG proteins in vitro.[88] Based on data from this laboratory, we speculate that the cleavage reactions are mediated in trans by a RAG1/RAG2 heterotetramer, but acknowledge that this organization remains to be fully validated. Genetic and biochemical evidence reviewed elsewhere[6,89] suggests that after cleavage, the signal and coding ends are held transiently in a four-end "post-cleavage synaptic complex", but coding ends are poorly retained within this complex, whereas the RAG proteins remain stably bound to the signal ends. This differential retention is reflected in the apparent uncoupling of coding and signal joint formation, with the former occurring more rapidly than the latter.

Transcription Factor-Assisted Targeting of Antigen Receptor Loci

Figure 2 presents a picture of the RAG proteins (with HMGB1/2) as being solely responsible for mediating synapsis during V(D)J recombination. However, this view is overly simplistic, because accumulating evidence suggests that the RAG complex can be preferentially targeted to specific antigen receptor loci through interactions with cellular factors that mark accessible and actively rearranging loci (such as modified histones),[90-92] or can bind to specific sites within particular antigen receptor genes.[93,94] Here we will briefly review studies of the latter class of RAG interaction factors and discuss the findings as they relate to RAG-RSS complex assembly.

Pax5 is a B lineage-specific transcription factor that regulates many B lineage-specific genes and is required to support rearrangement of D_H-distal V_H gene segments during lymphocyte development.[95] Zhang et al recently discovered that 94% of V_H coding regions (which are all flanked by a 23-RSS) contain two or more potential Pax5 binding sites.[93] The authors showed that Pax5 can indeed bind these sites and promote RAG-mediated cleavage and rearrangement of different V_H 23-RSSs when Pax5 binding sites are present in the flanking coding sequence. The authors further demonstrated that Pax5 directly interacts with the RAG proteins; this association requires the N-terminal paired DNA binding domain of Pax5 and is observed only when both RAG proteins are present. Based on these data, the authors conclude that Pax5 promotes V_H-to-DJ_H rearrangement by stabilizing RAG binding to the V_H 23-RSS via bridging interactions between the RAG proteins and the Pax5 binding site. Whether Pax5 binds the coding region first and then recruits the RAG complex to the 23-RSS, or, alternatively, whether Pax5 stably interacts with the RAG proteins before RSS engagement and maintains this association after the RAG proteins bind a 12-RSS in order to facilitate synapsis with a 23-RSS (containing Pax5 binding sites in the coding sequence) was not directly tested in this study. If the latter were true, one might expect that Pax5 could supershift a 12-RSS-RAG complex by EMSA.

How the established ordering of TCRβ locus rearrangements ($D_β$-to-$J_β$ recombination preceding $V_β$-to-$DJ_β$ rearrangement) is enforced remains in question.[96] To explain this phenomenon, Wang et al[94] investigated whether $D_β$ 23-RSSs contain a transcription factor recognition site(s) through which its binding could direct RAG-mediated $D_β$-to-$J_β$ rearrangement in preference to $V_β$-to-$DJ_β$ recombination. The authors provide evidence that TCR 3′-Dβ 23-RSSs contain an AP1 transcription factor binding site, which extends from the 6th bp of the heptamer to the 5th bp of the spacer and that the AP1 component c-Fos can bind to this sequence. c-Fos was shown to promote RAG association with a 3′ $D_β$ 23-RSS and enhance $D_β$-$J_β$ recombination in cells, while, conversely, reducing $V_β$-$D_β$ rearrangement. These effects were abolished if the putative c-Fos binding site was mutated. Mice deficient in c-Fos were shown to exhibit impaired TCRβ rearrangement overall, but elevated levels of mis-ordered $V_β$-$DJ_β$ recombination. Whether direct $V_β$-to-$J_β$ recombination was also elevated in these mice was not directly tested, but would have been interesting to determine because this rearrangement is formally permitted by the 12/23 rule. The authors showed that c-Fos associates with the core RAG proteins (primarily core RAG2), requiring the DNA binding domain and leucine zipper motif of c-Fos for this interaction. Interestingly, unlike Pax5,[93] the transcription-activation domain of c-Fos is not required to stimulate V(D)J rearrangement.[94] Thus, the authors conclude that c-Fos may facilitate the selective recruitment of the RAG proteins to the 3′$D_β$ 23-RSS, thereby promoting preferential $D_β$-$J_β$ rearrangement. As is the case for Pax5, the order of events that leads to c-Fos association with the RAG synaptic complex remains unclear. What is striking about the location of the AP1 binding site in the 3′$D_β$ 23-RSS is that it encompasses the same region contacted by the RAG proteins in a 23-RSS-RAG protein complex assembled in the presence of HMGB1.[51] Indeed, structural studies of AP1-DNA complexes[97] suggest that AP1 would engage this sequence in a manner similar to the RAG proteins,[24] interacting primarily with the major groove and contacting some of the same phosphodiester bonds in the RSS. Since the two protein complexes cannot occupy the same space, we speculate that in these complexes, RAG-mediated interactions with the RSS at this location are functionally replaced by AP1 contacts. The portions of the RAG proteins normally mediating these contacts may be freed to engage another DNA sequence. One intriguing possibility is that the displaced RAG DNA

binding domains contact the 3′ Dβ-12-RSS and through this engagement, help prevent it from becoming a target for synapsis with an upstream Vβ-23-RSS.

Conclusion and Future Directions

Accumulating evidence supports a capture model of RAG synaptic complex assembly and coupled RSS cleavage that is initiated by RAG binding and nicking of a 12-RSS and followed by the 23-RSS capture and cleavage of both RSSs using a base-flipping mechanism to facilitate hairpin formation. The stoichiometry and organization of the RAG proteins in the synaptic complex is still controversial and uncertain and will not likely be resolved until it yields to structural characterization. The base-flipping strategy used by the RAG proteins to mediate hairpin formation is also used by the Tn5 transposase during transposition and represents yet another parallel among the many mechanistic similarities between V(D)J recombination and transposition that have been recognized over the years.[98] There is little doubt that as years progress, additional features held in common between these systems will be discovered. One of the more difficult processes to understand in these systems is how synapsis is sensed. For the RAG proteins, this process remains mysterious, but the active site itself appears to play an important role, as mutations in RAG1 near the DDE motif have recently been identified that enable the RAG complex to mediate base-flipping and V(D)J rearrangement in violation of the 12/23 rule. The molecular basis for these effects remains to be elucidated. Recent evidence also suggests that the choice of which RSSs to assemble into a synaptic complex may be guided by interactions between the RAG proteins and other DNA binding factors. The finding that core RAG proteins interact with HMGB1/2 and, more recently, two different transcription factors, suggests that the core RAG1/2 complex contains one or more protein interaction domains potentially competent to mediate association with a variety of DNA binding proteins. This raises the possibility that previously observed differences in antigen receptor gene usage[99,100] may in some cases be explained by cellular factors that bind DNA at sites proximal to the RSS and promote RAG-RSS complex formation by direct interaction with the RAG proteins.

Acknowledgements

The authors wish to acknowledge support from the National Institutes of Health (R01 AI055599) and funding from the American Cancer Society (RSG-01-020-01-CCE).

References

1. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell 2002; 109(Suppl):S45-55.
2. Ramsden DA, Baetz K, Wu GE. Conservation of sequence in recombination signal sequence spacers. Nucleic Acids Res 1994; 22(10):1785-1796.
3. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell 1989; 59(6):1035-1048.
4. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248(4962):1517-1523.
5. Gellert M. V(D)J recombination: RAG proteins, repair factors and regulation. Annu Rev Biochem 2002; 71:101-132.
6. Fugmann SD, Lee AI, Shockett PE et al. The RAG proteins and V(D)J recombination: complexes, ends and transposition. Annu Rev Immunol 2000; 18:495-527.
7. Schlissel M, Constantinescu A, Morrow T et al. Double-strand signal sequence breaks in V(D)J recombination are blunt, 5′-phosphorylated, RAG-dependent and cell cycle regulated. Genes Dev 1993; 7(12B):2520-2532.
8. Roth DB, Menetski JP, Nakajima PB et al. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell 1992; 70(6):983-991.
9. McBlane JF, van Gent DC, Ramsden DA et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 1995; 83(3):387-395.
10. Lewis SM, Hesse JE, Mizuuchi K et al. Novel strand exchanges in V(D)J recombination. Cell 1988; 55(6):1099-1107.
11. Morzycka-Wroblewska E, Lee FE, Desiderio SV. Unusual immunoglobulin gene rearrangement leads to replacement of recombinational signal sequences. Science 1988; 242(4876):261-263.

12. Lieber MR, Ma Y, Pannicke U et al. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. DNA Repair (Amst) 2004; 3(8-9):817-826.
13. Buck D, Malivert L, de Chasseval R et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 2006; 124(2):287-299.
14. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 2006; 124(2):301-313.
15. Ma Y, Pannicke U, Schwarz K et al. Hairpin opening and overhang processing by an artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell 2002; 108(6):781-794.
16. Bertocci B, De Smet A, Berek C et al. Immunoglobulin kappa light chain gene rearrangement is impaired in mice deficient for DNA polymerase mu. Immunity 2003; 19(2):203-211.
17. Bertocci B, De Smet A, Weill JC et al. Nonoverlapping functions of DNA polymerases mu, lambda and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo. Immunity 2006; 25(1):31-41.
18. Weterings E, Chen DJ. The endless tale of nonhomologous end-joining. Cell Res 2008; 18(1):114-124.
19. Rooney S, Chaudhuri J, Alt FW. The role of the nonhomologous end-joining pathway in lymphocyte development. Immunol Rev 2004; 200:115-131.
20. Lieber MR, Lu H, Gu J et al. Flexibility in the order of action and in the enzymology of the nuclease, polymerases and ligase of vertebrate nonhomologous DNA end joining: relevance to cancer, aging and the immune system. Cell Res 2008; 18(1):125-133.
21. De P, Rodgers KK. Putting the pieces together: identification and characterization of structural domains in the V(D)J recombination protein RAG1. Immunol Rev 2004; 200:70-82.
22. Eastman QM, Leu TM, Schatz DG. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature 1996; 380(6569):85-88.
23. van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell 1996; 85(1):107-113.
24. Swanson PC. The bounty of RAGs: recombination signal complexes and reaction outcomes. Immunol Rev 2004; 200:90-114.
25. Difilippantonio MJ, McMahan CJ, Eastman QM et al. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell 1996; 87(2):253-262.
26. Spanopoulou E, Zaitseva F, Wang FH et al. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. Cell 1996; 87(2):263-276.
27. Swanson PC, Desiderio S. RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. Mol Cell Biol 1999; 19(5):3674-3683.
28. Ciubotaru M, Ptaszek LM, Baker GA et al. RAG1-DNA binding in V(D)J recombination. Specificity and DNA-induced conformational changes revealed by fluorescence and CD spectroscopy. J Biol Chem 2003; 278(8):5584-5596.
29. Bailin T, Mo X, Sadofsky MJ. A RAG1 and RAG2 tetramer complex is active in cleavage in V(D)J recombination. Mol Cell Biol 1999; 19(7):4664-4671.
30. Rodgers KK, Villey IJ, Ptaszek L et al. A dimer of the lymphoid protein RAG1 recognizes the recombination signal sequence and the complex stably incorporates the high mobility group protein HMG2. Nucleic Acids Res 1999; 27(14):2938-2946.
31. Godderz LJ, Rahman NS, Risinger GM et al. Self-association and conformational properties of RAG1: implications for formation of the V(D)J recombinase. Nucleic Acids Res 2003; 31(7):2014-2023.
32. Swanson PC, Desiderio S. V(D)J recombination signal recognition: distinct, overlapping DNA-protein contacts in complexes containing RAG1 with and without RAG2. Immunity 1998; 9(1):115-125.
33. Hiom K, Gellert M. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell 1997; 88(1):65-72.
34. Mundy CL, Patenge N, Matthews AG et al. Assembly of the RAG1/RAG2 synaptic complex. Mol Cell Biol 2002; 22(1):69-77.
35. Santagata S, Aidinis V, Spanopoulou E. The effect of Me2+ cofactors at the initial stages of V(D)J recombination. J Biol Chem 1998; 273(26):16325-16331.
36. Mo X, Bailin T, Sadofsky MJ. RAG1 and RAG2 cooperate in specific binding to the recombination signal sequence in vitro. J Biol Chem 1999; 274(11):7025-7031.
37. Swanson PC. A RAG-1/RAG-2 tetramer supports 12/23-regulated synapsis, cleavage and transposition of V(D)J recombination signals. Mol Cell Biol 2002; 22(22):7790-7801.
38. Raval P, Kriatchko AN, Kumar S et al. Evidence for Ku70/Ku80 association with full-length RAG1. Nucleic Acids Res 2008; 36(6):2060-2072.
39. De P, Zhao S, Gwyn LM et al. Thermal dependency of RAG1 self-association properties. BMC Biochem 2008; 9:5.

40. Kim DR, Dai Y, Mundy CL et al. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. Genes Dev 1999; 13(23):3070-3080.
41. Landree MA, Wibbenmeyer JA, Roth DB. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. Genes Dev 1999; 13(23):3059-3069.
42. Fugmann SD, Villey IJ, Ptaszek LM et al. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. Mol Cell 2000; 5(1):97-107.
43. Haren L, Ton-Hoang B, Chandler M. Integrating DNA: transposases and retroviral integrases. Annu Rev Microbiol 1999; 53:245-281.
44. Davies DR, Goryshin IY, Reznikoff WS et al. Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science 2000; 289(5476):77-85.
45. Naumann TA, Reznikoff WS. Trans catalysis in Tn5 transposition. Proc Natl Acad Sci USA 2000; 97(16):8944-8949.
46. Landree MA, Kale SB, Roth DB. Functional organization of single and paired V(D)J cleavage complexes. Mol Cell Biol 2001; 21(13):4256-4264.
47. Swanson PC. The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. Mol Cell Biol 2001; 21(2):449-458.
48. van Gent DC, Hiom K, Paull TT et al. Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J 1997; 16(10):2665-2670.
49. Sawchuk DJ, Weis-Garcia F, Malik S et al. V(D)J recombination: modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. J Exp Med 1997; 185(11):2025-2032.
50. Calogero S, Grassi F, Aguzzi A et al. The lack of chromosomal protein hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. Nat Genet 1999; 22(3):276-280.
51. Swanson PC. Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. Mol Cell Biol 2002; 22(5):1340-1351.
52. Thomas JO, Travers AA. HMG1 and 2 and related 'architectural' DNA-binding proteins. Trends Biochem Sci 2001; 26(3):167-174.
53. Klune JR, Dhupar R, Cardinal J et al. Hmgb1: Endogenous danger signaling. Mol Med 2008; 14(7-8):476-484.
54. Aidinis V, Bonaldi T, Beltrame M et al. The RAG1 homeodomain recruits HMG1 and HMG2 to facilitate recombination signal sequence binding and to enhance the intrinsic DNA-bending activity of RAG1-RAG2. Mol Cell Biol 1999; 19(10):6532-6542.
55. Yoshida T, Tsuboi A, Ishiguro K et al. The DNA-bending protein, HMG1, is required for correct cleavage of 23 bp recombination signal sequences by recombination activating gene proteins in vitro. Int Immunol 2000; 12(5):721-729.
56. Bergeron S, Madathiparambil T, Swanson PC. Both high mobility group (HMG)-boxes and the acidic tail of HMGB1 regulate recombination-activating gene (RAG)-mediated recombination signal synapsis and cleavage in vitro. J Biol Chem 2005; 280(35):31314-31324.
57. Kriatchko AN, Bergeron S, Swanson PC. HMG-box domain stimulation of RAG1/2 cleavage activity is metal ion dependent. BMC Mol Biol 2008; 9:32.
58. Dai Y, Wong B, Yen YM et al. Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination. Mol Cell Biol 2005; 25(11):4413-4425.
59. Bonaldi T, Langst G, Strohner R et al. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J 2002; 21(24):6865-6873.
60. Ueda T, Chou H, Kawase T et al. Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription stimulation. Biochemistry 2004; 43(30):9901-9908.
61. Kwon J, Imbalzano AN, Matthews A et al. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. Mol Cell 1998; 2(6):829-839.
62. Nightingale KP, Baumann M, Eberharter A et al. Acetylation increases access of remodelling complexes to their nucleosome targets to enhance initiation of V(D)J recombination. Nucleic Acids Res 2007; 35(18):6311-6321.
63. Baumann M, Mamais A, McBlane F et al. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. EMBO J 2003; 22(19):5197-5207.
64. Jones JM, Gellert M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. EMBO J 2002; 21(15):4162-4171.
65. Curry JD, Geier JK, Schlissel MS. Single-strand recombination signal sequence nicks in vivo: evidence for a capture model of synapsis. Nat Immunol 2005; 6(12):1272-1279.
66. Eastman QM, Schatz DG. Nicking is asynchronous and stimulated by synapsis in 12/23 rule-regulated V(D)J cleavage. Nucleic Acids Res 1997; 25(21):4370-4378.

67. Yu K, Lieber MR. The nicking step in V(D)J recombination is independent of synapsis: implications for the immune repertoire. Mol Cell Biol 2000; 20(21):7914-7921.
68. Sheehan KM, Lieber MR. V(D)J recombination: signal and coding joint resolution are uncoupled and depend on parallel synapsis of the sites. Mol Cell Biol 1993; 13(3):1363-1370.
69. Ciubotaru M, Kriatchko AN, Swanson PC et al. Fluorescence resonance energy transfer analysis of recombination signal sequence configuration in the RAG1/2 synaptic complex. Mol Cell Biol 2007; 27(13):4745-4758.
70. Nagawa F, Ishiguro K, Tsuboi A et al. Footprint analysis of the RAG protein recombination signal sequence complex for V(D)J type recombination. Mol Cell Biol 1998; 18(1):655-663.
71. Akamatsu Y, Oettinger MA. Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. Mol Cell Biol 1998; 18(8):4670-4678.
72. Eastman QM, Villey IJ, Schatz DG. Detection of RAG protein-V(D)J recombination signal interactions near the site of DNA cleavage by UV cross-linking. Mol Cell Biol 1999; 19(5):3788-3797.
73. Mo X, Bailin T, Noggle S et al. A highly ordered structure in V(D)J recombination cleavage complexes is facilitated by HMG1. Nucleic Acids Res 2000; 28(5):1228-1236.
74. Nagawa F, Kodama M, Nishihara T et al. Footprint analysis of recombination signal sequences in the 12/23 synaptic complex of V(D)J recombination. Mol Cell Biol 2002; 22(20):7217-7225.
75. Nagawa F, Hirose S, Nishizumi H et al. Joining mutants of RAG1 and RAG2 that demonstrate impaired interactions with the coding-end DNA. J Biol Chem 2004; 279(37):38360-38368.
76. Tsai CL, Drejer AH, Schatz DG. Evidence of a critical architectural function for the RAG proteins in end processing, protection and joining in V(D)J recombination. Genes Dev 2002; 16(15):1934-1949.
77. Qiu JX, Kale SB, Yarnell Schultz H et al. Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol Cell 2001; 7(1):77-87.
78. Cuomo CA, Mundy CL, Oettinger MA. DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. Mol Cell Biol 1996; 16(10):5683-5690.
79. Ramsden DA, McBlane JF, van Gent DC et al. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. EMBO J 1996; 15(12):3197-3206.
80. Grundy GJ, Hesse JE, Gellert M. Requirements for DNA hairpin formation by RAG1/2. Proc Natl Acad Sci USA 2007; 104(9):3078-3083.
81. Bhasin A, Goryshin IY, Reznikoff WS. Hairpin formation in Tn5 transposition. J Biol Chem 1999; 274(52):37021-37029.
82. Nishihara T, Nagawa F, Imai T et al. RAG-heptamer interaction in the synaptic complex is a crucial biochemical checkpoint for the 12/23 recombination rule. J Biol Chem 2008; 283(8):4877-4885.
83. Ason B, Reznikoff WS. Mutational analysis of the base flipping event found in Tn5 transposition. J Biol Chem 2002; 277(13):11284-11291.
84. Bischerour J, Chalmers R. Base-flipping dynamics in a DNA hairpin processing reaction. Nucleic Acids Res 2007; 35(8):2584-2595.
85. Lu CP, Sandoval H, Brandt VL et al. Amino acid residues in rag1 crucial for DNA hairpin formation. Nat Struct Mol Biol 2006; 13(11):1010-1015.
86. Bergeron S, Anderson DK, Swanson PC. RAG and HMGB1 proteins: purification and biochemical analysis of recombination signal complexes. Methods Enzymol 2006; 408:511-528.
87. Kriatchko AN, Anderson DK, Swanson PC. Identification and characterization of a gain-of-function RAG-1 mutant. Mol Cell Biol 2006; 26(12):4712-4728.
88. Ciubotaru M, Schatz DG. Synapsis of recombination signal sequences located in cis and DNA underwinding in V(D)J recombination. Mol Cell Biol 2004; 24(19):8727-8744.
89. Schatz DG. V(D)J recombination moves in vitro. Semin Immunol 1997; 9(3):149-159.
90. West KL, Singha NC, De Ioannes P et al. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. Immunity 2005; 23(2):203-212.
91. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27(4):561-571.
92. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450(7172):1106-1110.
93. Zhang Z, Espinoza CR, Yu Z et al. Transcription factor pax5 (BSAP) transactivates the RAG-mediated V(H)-to-DJ(H) rearrangement of immunoglobulin genes. Nat Immunol 2006; 7(6):616-624.
94. Wang X, Xiao G, Zhang Y et al. Regulation of tcrb recombination ordering by c-fos-dependent RAG deposition. Nat Immunol 2008; 9(7):794-801.
95. Cobaleda C, Schebesta A, Delogu A et al. Pax5: the guardian of B-cell identity and function. Nat Immunol 2007; 8(5):463-470.
96. Jackson AM, Krangel MS. Turning T-cell receptor beta recombination on and off: more questions than answers. Immunol Rev 2006; 209:129-141.

97. Chen L, Glover JN, Hogan PG et al. Structure of the DNA-binding domains from NFAT, fos and jun bound specifically to DNA. Nature 1998; 392(6671):42-48.
98. Jones JM, Gellert M. The taming of a transposon: V(D)J recombination and the immune system. Immunol Rev 2004; 200:233-248.
99. Feeney AJ, Tang A, Ogwaro KM. B-cell repertoire formation: role of the recombination signal sequence in nonrandom V segment utilization. Immunol Rev 2000; 175:59-69.
100. Livak F, Burtrum DB, Rowen L et al. Genetic modulation of T-cell receptor gene segment usage during somatic recombination. J Exp Med 2000; 192(8):1191-1196.
101. Ma Y, Lu H, Tippin B et al. A biochemically defined system for mammalian nonhomologous DNA end joining. Mol Cell 2004; 16(5):701-713.

Chapter 2

Regulation of RAG Transposition

Adam G.W. Matthews* and Marjorie A. Oettinger

Abstract

V(D)J recombination is initiated by the lymphoid specific proteins RAG1 and RAG2, which together constitute the V(D)J recombinase. However, the RAG1/2 complex can also act as a transposase, inserting the broken DNA molecules generated during V(D)J recombination into an unrelated piece of DNA. This process, termed RAG transposition, can potentially cause insertional mutagenesis, chromosomal translocations and genomic instability. This review focuses on the mechanism and regulation of RAG transposition. We first provide a brief overview of the biochemistry of V(D)J recombination. We then discuss the discovery of RAG transposition and present an overview of the RAG transposition pathway. Using this pathway as a framework, we discuss the factors and forces that regulate RAG transposition.

Introduction

During lymphoid development, immunoglobulin and T-cell receptor genes are assembled from multiple, nonconsecutive gene segments in a series of site-specific recombination reactions, termed V(D)J recombination.[1,2] By combinatorially joining different variable (V), diversity (D) and joining (J) gene segments, V(D)J recombination generates a diverse array of T-cell receptor (TCR) and immunoglobulin (Ig) molecules (Fig. 1), thereby enabling the adaptive immune system to recognize an almost limitless number of antigens and protect us from pathogenic microorganisms.

V(D)J recombination is initiated when the lymphoid specific proteins RAG1 and RAG2 generate double-stranded DNA breaks at V, D and J gene segments. These breaks are normally repaired by the nonhomologous end-joining (NHEJ) pathway. However, the same enzyme that produces these double-strand breaks—RAG1/2 complex—can also act as a transposase, inserting the newly generated broken DNA molecules into an unrelated piece of DNA. This process, termed RAG transposition, can not only cause insertional mutagenesis,[3] but could also lead to genomic instability[4] and the generation of potentially oncogenic chromosomal translocations.[5] Therefore, it is important to understand how RAG transposition is suppressed in vivo.

This review will focus on the mechanism and regulation of RAG transposition. We will first provide a brief overview of the biochemistry of V(D)J recombination. We will then discuss the discovery of RAG transposition and present an overview of the RAG transposition pathway. Using this pathway as a framework, the factors and forces that regulate RAG transposition will be discussed.

Biochemistry of V(D)J Recombination

All recombinationally active V, D and J gene segments are flanked by recombination signal sequences (RSSs),[6] which consist of highly conserved heptamer (5′-CACAGTG-3′) and nonamer (5′-ACAAAAACC-3′) sequences separated by a spacer region of either 12 or 23 bp.[7-10] Efficient

*Corresponding Author: Adam G.W. Matthews—Department of Molecular Biology, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, MA 02114, USA. Email: agwmatthews@gmail.com

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

recombination between gene segments only occurs when one segment is flanked by a 12-RSS and the other is flanked by a 23-RSS, a restriction termed the 12/23 rule.[6] The 12/23 rule ensures that recombination only occurs between gene segments that can give rise to a functional antigen receptor gene.

V(D)J recombination requires the expression of two lymphoid-specific recombination-activating genes, RAG1 and RAG2,[11-15] which act together to constitute the V(D)J recombinase that recognizes and cleaves recombination signal sequences.[16] Although RAG1 and RAG2 can both be truncated down to catalytically active "core" regions, consisting of amino acids 384-1008 out of 1040 for RAG1[17-19] and amino acids 1-383 out of 527 for RAG2,[20,21] the "non-core" portions of RAG1 and RAG2, which are highly conserved throughout evolution,[22,23] play key regulatory roles in vivo.[24-29]

V(D)J recombination can be conceptually divided into two stages: generation of double-stranded DNA breaks by the lymphoid-specific proteins RAG1 and RAG2[16] and the repair of those breaks by nonhomologous end-joining. While DNA double-strand break formation (V(D)J cleavage) requires only the RAG proteins and HMG1 (a DNA bending protein), the repair stage of the reaction requires the ubiquitously expressed nonhomologous end-joining (NHEJ) proteins Ku70,[30,31] Ku80,[32,33] DNA-PK_{CS},[34,35] Artemis,[36,37] XRCC4,[38] DNA Ligase IV[39,40] and XLF (a.k.a. Cernunnos).[41-44] The RAG proteins also play a role in the repair stage of the reaction.[45-51] Additionally, other proteins such as ATM, Mre11, Rad50 and Nbs1 may also be involved in the repair of RAG-induced double-strand breaks.[52,53]

During the cleavage stage of the reaction, the RAG1/2 complex first assembles on a 12-RSS and then captures a 23-RSS[54-56] to form a synaptic paired complex.[57] DNA double-strand breaks

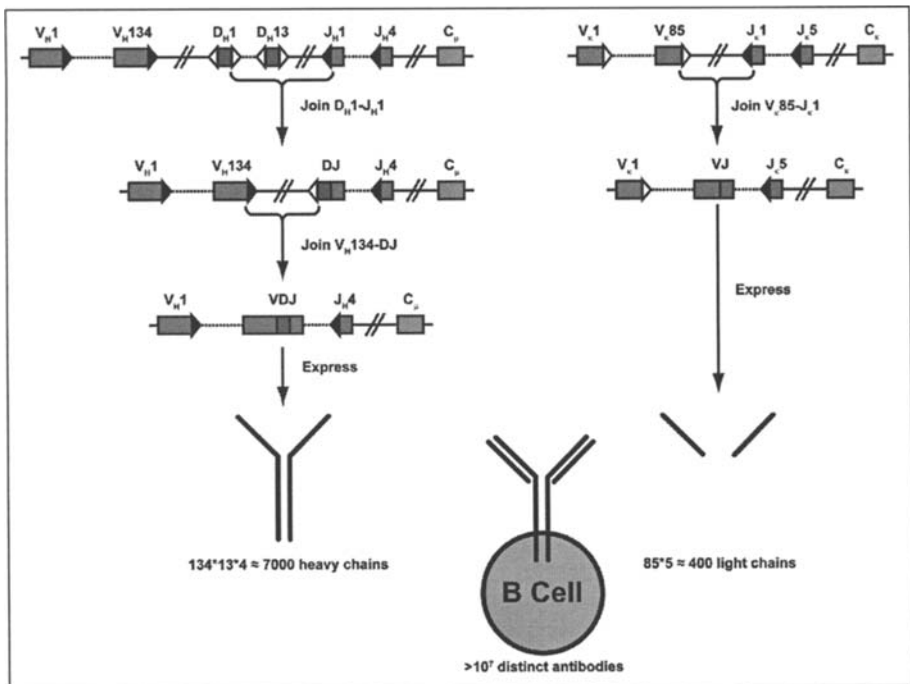


Figure 1. Overview of V(D)J recombination. V, D, and J gene segments are depicted as rectangles, constant region genes are depicted as rectangles and recombination signal sequences (RSSs) are depicted as triangles (shaded for the 23-RSS and unshaded for the 12-RSS). As a result of V(D)J recombination, our bodies generate a diverse repertoire of antigen receptors from a limited amount of genetic material. A color version of this figure is available online at www.landesbioscience.com/curie.

are generated within this paired complex via a pair of phosphoryl transfer reactions (Fig. 2, top). The RAG proteins first nick the top strand of each RSS, just 5′ of the heptamer sequence.[16] These newly liberated 3′ hydroxyl groups on the top strand of the coding flank then attack the bottom strand via a direct transesterification,[58] thereby converting these DNA nicks into double-strand breaks and generating the cleavage reaction products: two hairpinned coding ends and two blunt 5′-phosphorylated signal ends.[16]

During the repair phase of the reaction, Ku70/Ku80 heterodimers are thought to bind to the four cleavage products. DNA-PK_{CS} then binds to Artemis and undergoes autophosphorylation, thereby enabling Artemis to endonucleolytically open the hairpinned coding ends.[59,60] Since hairpin-opening rarely occurs precisely at the tip of the hairpin, 5′ or 3′ overhangs are commonly generated. These overhangs may then be trimmed by nucleases or filled in by polymerases, generating palindromic (P) nucleotides. The two processed coding ends are then ligated together in a process requiring the XRCC4-XLF-DNA Ligase IV complex to form the coding joint.[61-63] The XRCC4-XLF-DNA Ligase IV complex also repairs the two signal ends by precise heptamer-to-heptamer ligation to form the signal joint.

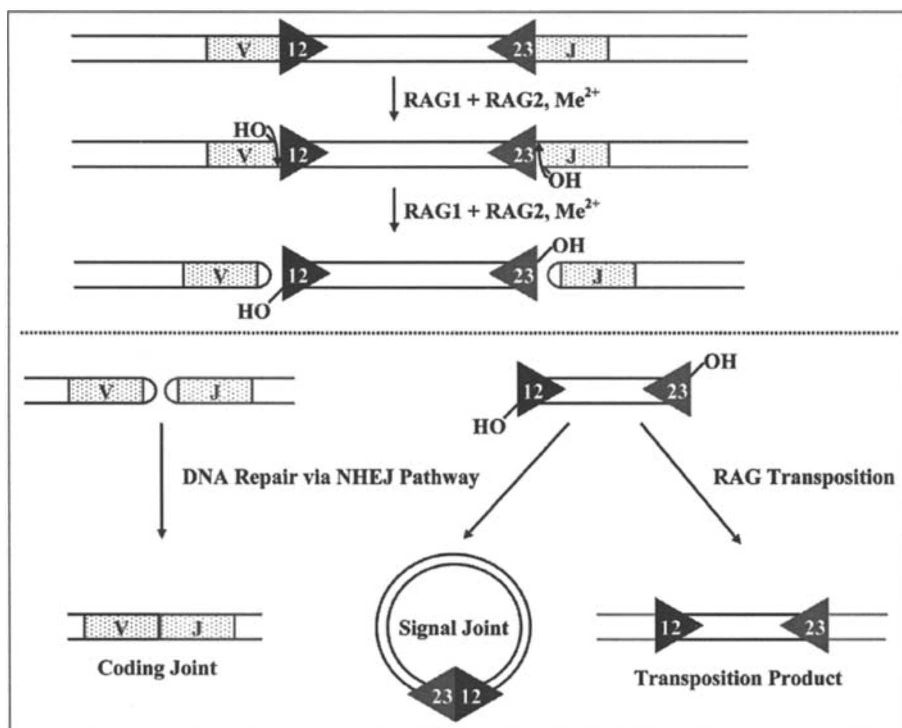


Figure 2. The biochemistry of V(D)J recombination and RAG transposition. V gene segments are depicted as stippled rectangles, J gene segments are depicted as stippled rectangles, recombination signal sequences are depicted as triangles (for the 12-RSS and for the 23-RSS) and target DNA is shown in light gray. The RAG1/2 complex initiates V(D)J recombination by first nicking DNA at the border between the coding DNA and the RSS heptamer (hydrolysis). The free 3′ hydroxyl (OH) on the coding flank then attacks the opposite strand in a direct transesterification to form a blunt signal end and a hairpinned coding end. Hairpinned coding ends are repaired via the nonhomologous end-joining (NHEJ) pathway to form imprecise coding joints while blunt signal ends can either be repaired via the NHEJ pathway to form precise signal joints, or they can inserted into an unrelated piece of DNA via the RAG transposition pathway. A color version of this figure is available online at www.landesbioscience.com/curie.

Overview of RAG Transposition

Discovery of RAG Transposition

Based on the palindromic sequence of the RSS's heptamer,[64] the fact that the genomic orientation of Igκ recombination signal sequences resembles the inverted repeats found at the ends of prokaryotic transposons[64] and the unusual structure of the mammalian RAG locus (RAG1 and RAG2 are compactly organized as adjacent genes and each gene is encoded by a single exon),[11] it was hypothesized that V(D)J recombination may be mechanistically related to bacterial transposition events.[11,64,65] And indeed, 10 years ago, it was shown that the RAG proteins can transpose signal ends into an unrelated piece of DNA, in a process termed RAG transposition (Fig. 2, bottom).[5,66] In this reaction, the RAG proteins catalyze another phosphoryl transfer reaction, enabling the exposed, nucleophilic 3′ hydroxyl group on the bottom strand of the signal end to attack a target DNA molecule. Although RAG transposition was initially discovered in vitro, rare events of RAG transposition have also been observed in vivo in human cells,[3,67] murine cells[4] and in engineered yeast.[68] Thus, RAG transposition represents a bona fide alternative fate for the double-strand breaks generated during V(D)J recombination. By competing with the NHEJ pathway in vivo, transposition can cause insertional mutagenesis,[3] oncogenic chromosomal translocations[5] and genomic instability.[4] The pathways leading to insertional mutagenesis and chromosomal translocations are described in more detail below.

The RAG Transposition Pathway

As diagrammed in Figure 3, RAG transposition proceeds through an orderly series of steps.[69] The RAG proteins first bind to both a 12-RSS and a 23-RSS to form a synaptic paired complex (PC). The RAG proteins then perform coupled cleavage to generate a pair of DNA double-strand breaks (Fig. 3, Step 1), resulting in the cleaved signal complex (CSC) which contains two blunt 3′-hydroxylated signal ends and two hairpinned coding ends. Next, coding ends are transferred from the cleaved signal complex to the NHEJ pathway, leaving the RAG proteins bound to signal ends within the signal-end complex (SEC) (Fig. 3, Step 2). The decision to resolve signal ends via NHEJ (Fig. 3, Step 3a) or RAG transposition occurs within the signal-end complex. If the RAG proteins bind target DNA and commit to undergoing transposition (Fig. 3, Step 3b), they first form a target capture complex (TCC). Within the target capture complex, the RAG proteins can catalyze either single-ended insertion of just one RSS (Fig. 3, Step 4a), or double-ended insertion of both RSSs (Fig. 3, Step 4b).

Resolution of Branched Transposition Intermediates Can Lead to Either Insertional Mutagenesis or Chromosomal Translocations

After double-ended insertion, the resulting branched DNA molecule can be resolved in one of three ways. It can be resolved by DNA repair, resulting in insertional mutagenesis with the characteristic 5 bp target site duplication (Fig. 4a).[5] Alternatively, the branched transposition intermediate can be resolved via disintegration. In this RAG-catalyzed reaction, the nucleophilic 3′ hydroxyls on the target DNA attack the newly formed phosphodiester bonds linking the RSSs to the target DNA, thereby regenerating both the cleaved signal end and the target DNA (Fig. 4b; Fig. 3, Steps 4b and 4d).[70] Finally, the branched molecule can be resolved via RAG-catalyzed hairpin formation. In this reaction, which is analogous to the formation of hairpinned coding ends during V(D)J cleavage, the nucleophilic 3′ hydroxyls on the target DNA attack the phosphodiester bond on the opposite strand of the target molecule, thereby generating hairpinned target ends and signal ends with 3′ overhangs of 5 nucleotides (Fig. 4c).[70] Joining these hairpinned target ends to the hairpinned coding ends would lead to reciprocal chromosomal translocations that could be potentially oncogenic (Fig. 5).[5,70] It is worth noting that translocations generated in this manner would not bear the hallmark target site duplications that are characteristic of traditional RAG transposition (resulting in insertional mutagenesis), even though they were generated as a result of RAG-catalyzed transposition and hairpin formation.

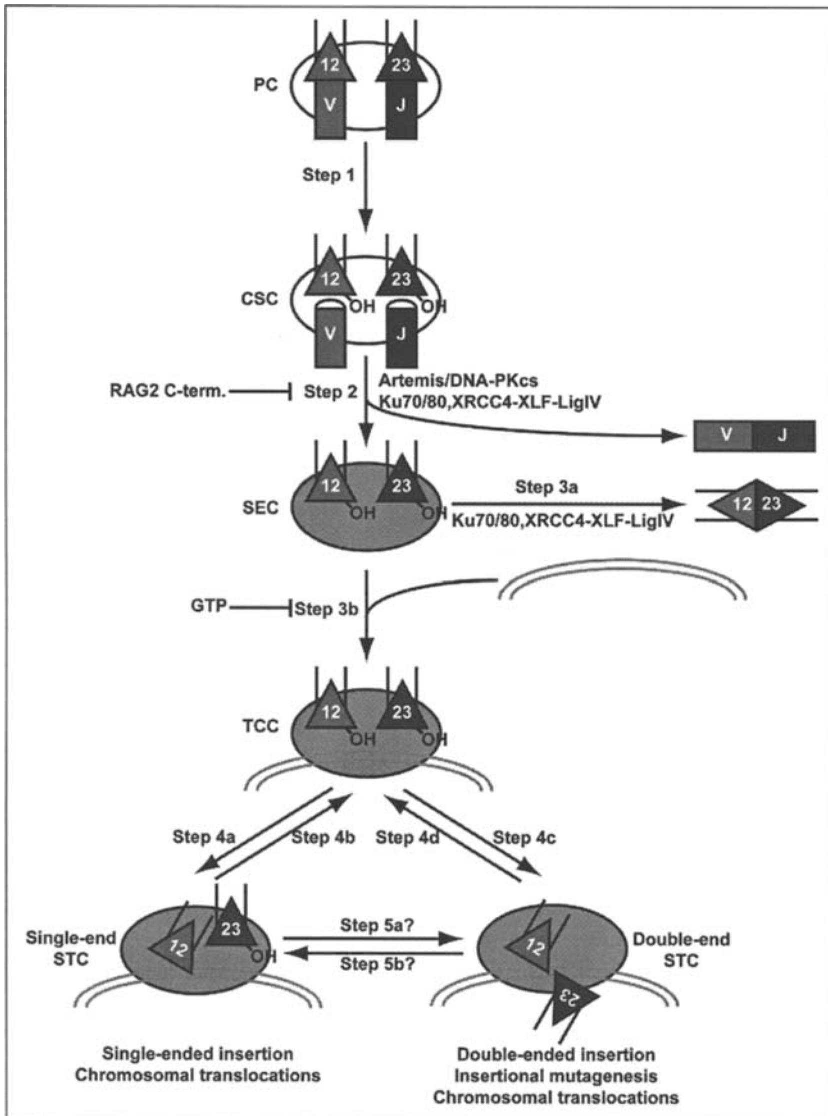


Figure 3. The RAG transposition pathway. V gene segments are depicted as rectangles, J gene segments are depicted as rectangles, recombination signal sequences are depicted as triangles (for the 12-RSS and for the 23-RSS), the RAG1/2 complex is portrayed as a shaded oval (either shaded or unshaded) and target DNA is shown in light gray. As described in the text, RAG transposition initiates with coupled RSS cleavage within the paired complex (PC), thereby generating the cleaved signal complex (CSC), which consists of all 4 broken DNA ends noncovalently bound by the RAG1/2 complex (Step 1). Transfer of the cleaved coding ends to the NHEJ pathway results in the formation of the signal end complex (SEC) (Step 2). Target capture occurs within the signal end complex, leading to the formation of a stable target capture complex (TCC) (Step 3b). Transpositional strand-transfer occurs within the target capture complex, generating the strand transfer complex (STC) (Steps 4a/4c). The branched DNA molecules present in the strand transfer complex can be resolved in several different ways (see Fig. 4). A color version of this figure is available online at www.landesbioscience.com/curie.

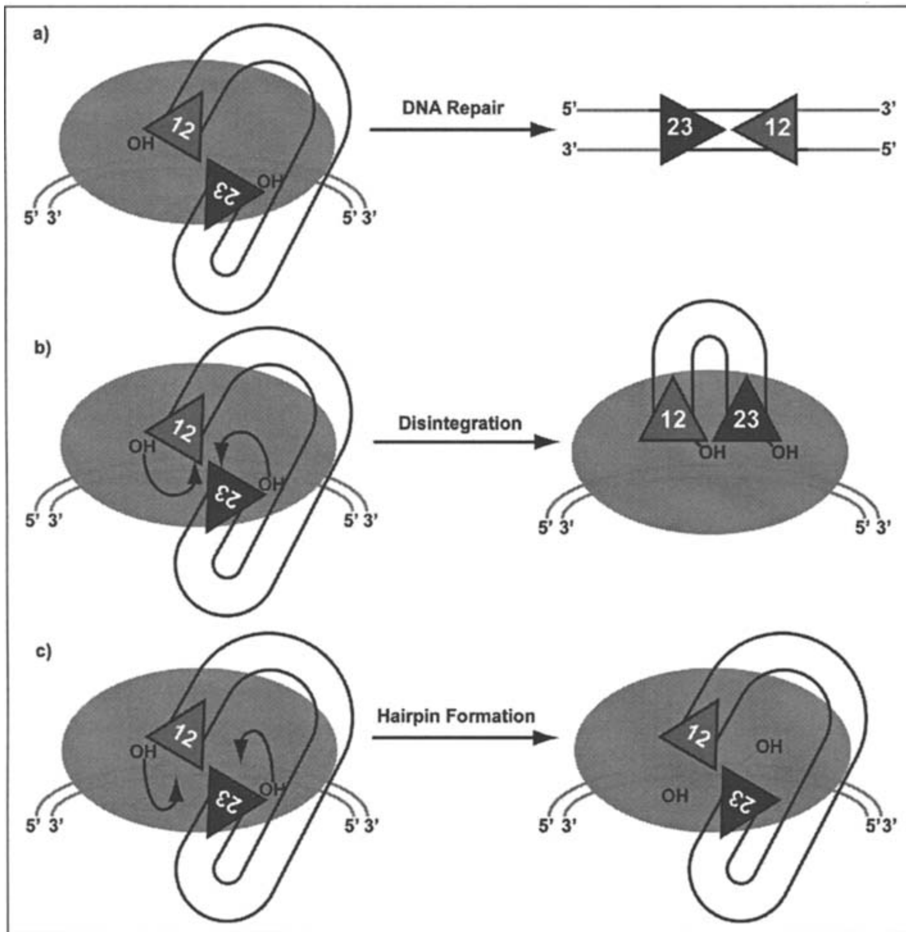


Figure 4. Pathways for resolving branched RAG transposition intermediates. After double-ended insertion (Fig. 3, Step 4c), the RAG1/2 complex remains bound to the branched transposition intermediate within the strand transfer complex. There are at least three pathways for resolving these transposition intermediates. a) These branched molecules can be resolved by nick repair, leading to insertional mutagenesis with the signal ends flanked by 5 bp target site duplications (shown in dark gray). b) The branched molecules can be resolved via disintegration, where the 3′ hydroxyls on the target DNA attack the phosphodiester bonds at the RSS-target DNA junctions, thereby removing the inserted signal ends and rejoining the target DNA. c) The branched molecules can be resolved via target cleavage, where the 3′ hydroxyls on the target DNA attack the opposite strand of the target DNA, thereby generating hairpinned target ends and liberating signal ends that contain 3′ overhangs of 5 nt. A color version of this figure is available online at www.landesbioscience.com/curie.

Single-ended insertion, followed by target DNA transesterification, can also lead to chromosomal translocations. If these translocations result from insertion of a signal end generated in a single-site cleavage event, they will be reciprocal.[5] If, however, they result from insertion of a signal end generated in a coupled cleavage event, they can be either reciprocal or nonreciprocal with loss of genetic material (Fig. 6). Nonreciprocal chromosomal translocations that are accompanied by loss of genetic material would likely lead to impaired viability of the cell and would therefore be difficult to detect in vivo. It

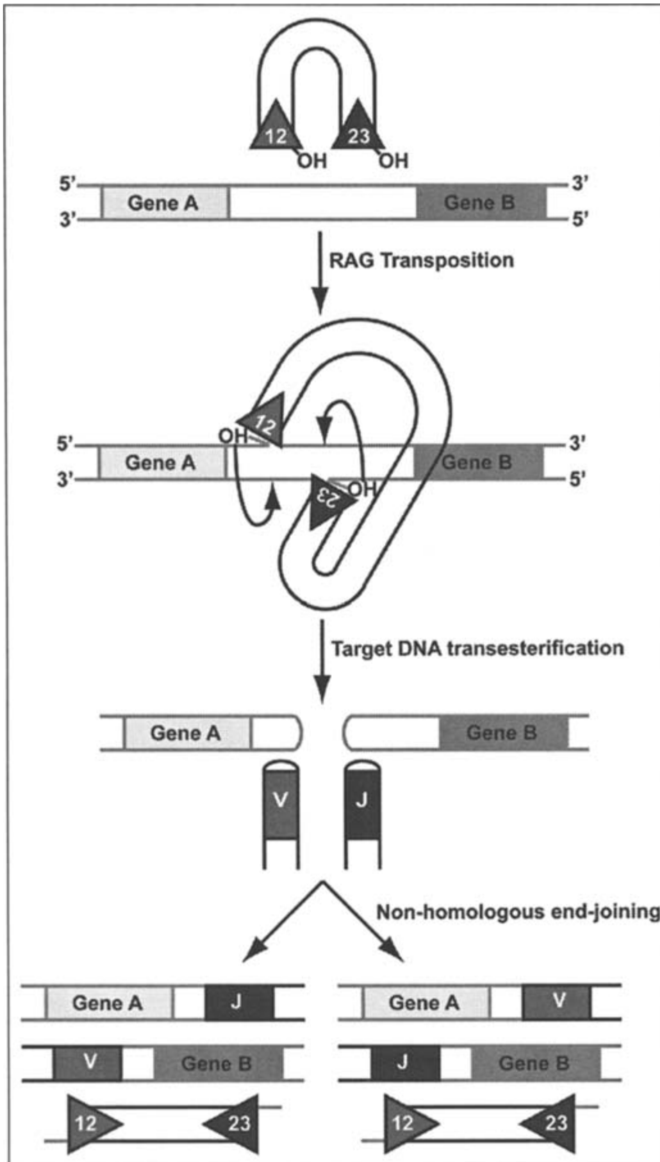


Figure 5. Pathway for generating chromosomal translocations from RAG transpositional double-ended insertions. Following RAG transpositional double-ended insertion, the branched transposition intermediate can be resolved via target DNA cleavage. Joining these hairpinned target ends to the previously generated hairpinned coding ends via nonhomologous end-joining would generate reciprocal translocations. If these translocations bring oncogenes (such as Gene A or Gene B) into close proximity with the immunoglobulin promoters/enhancers, they could potentially lead to oncogenic transformation of the cell. It is worth noting that in the same way that signal ends are normally lost as circular signal joints during canonical V(D)J recombination, the modified signal ends generated in the proceses of target DNA cleavage would also be lost either as linear DNA molecules (if left unprocessed) or as circular signal joints (if repaired by NHEJ).

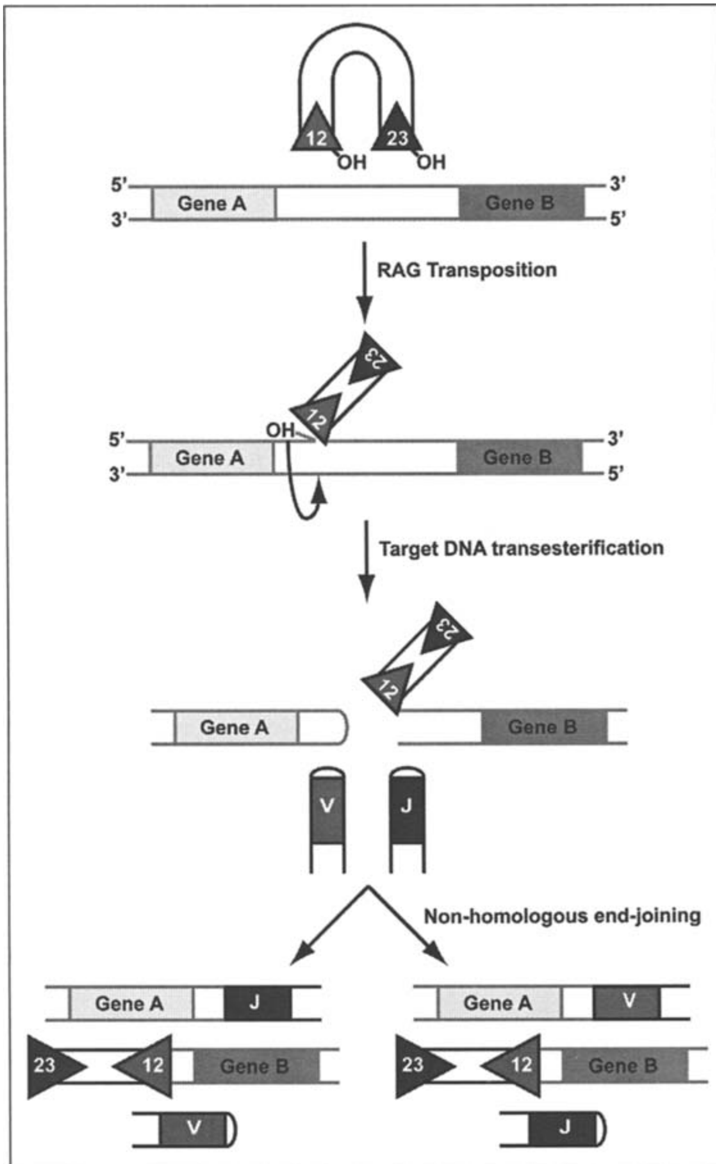


Figure 6. Pathway for generating chromosomal translocations from RAG transpositional single-ended insertions. Following RAG transpositional single-ended insertion, the branched transposition intermediate can be resolved via target DNA cleavage. Joining this hairpinned target end (in this case, Gene A) to one of the previously generated hairpinned coding ends via nonhomologous end-joining would generate a chromosomal translocation. Joining the remaining hairpinned coding end to the free signal end (in this case, 23-RSS) via either a RAG-dependent process (such as hybrid joining or open-and-shut joining) or nonhomologous end-joining would generate a reciprocal translocation. If, however, the remaining hairpinned coding end is not joined to the free signal end (as shown here), then this pathway would result in a nonreciprocal translocation with loss of genetic material, thereby impairing the viability of the cell.

is worth noting that although translocations resulting from single-ended insertion differ from those resulting from double-ended insertion, translocations generated as a result of single-ended insertion would also lack the target site duplications that are characteristic of traditional RAG transposition.

Regulation of RAG Transposition

As described above, RAG transposition events are potentially deleterious. Thus, it is vitally important that RAG transposition be suppressed in developing B-cells and T-cells. Indeed, although RAG transposition occurs robustly in vitro,[5,66] the frequency of transposition in vivo is much lower. One study estimated that in developing T-cells, RAG-mediated insertional mutagenesis causes phenotypic loss of HPRT gene function at a frequency of 1 event per 10^7 cells.[3] Since this study could not detect transposition events that occurred at other locations in the genome, the actual frequency of RAG transposition must be greater than 1 event per 10^7 cells. Another study estimated that in transfected 293T-cells, RAG transposition occurred at a frequency of 1 event per 10^7 plasmids analyzed.[67] A third study estimated that in pre-B-cells, one RAG transposition event occurs per every 50,000 V(D)J recombinations, corresponding to a frequency of 2.5 events per 10^5 recombinations.[4] Although all three studies conclude that RAG transposition occurs at a fairly low frequency in vivo, our bodies generate ~10^8 new lymphocytes per day. As such, the frequency of RAG transposition is high enough to cause potentially oncogenic genomic rearrangements. Therefore, RAG transposition is a biologically relevant pathway and it is important to understand the multiple ways in which it is regulated.

Using the pathway in Figure 3 as a framework, we can gain a better understanding of the mechanisms involved in the regulation of RAG transposition. In theory, RAG transposition could be suppressed at any of the four steps that precede donor insertion into the target DNA: RSS cleavage (Step 1); coding end release (Step 2); target capture (Step 3); or donor insertion (Step 4). However, whereas many transposons are regulated either at the level of transposase expression or at the step of transposon excision from the host genome (Fig. 3, Step 1), RAG transposition cannot be similarly regulated because RSS cleavage is crucial for the assembly of functional antigen receptor genes. Therefore, RAG-mediated transposition must be regulated at a step subsequent to RSS donor cleavage (Steps 2-4). Below, we review what is currently known about the factors and forces that regulate RAG transposition and we speculate about additional potential regulatory mechanisms.

Current Understanding of How RAG Transposition Is Regulated

Regulation by the C-Terminal Portion of RAG2

In vitro studies comparing the transpositional activity of full-length RAG2 (aa 1-527) to that of core RAG2 (aa 1-387) revealed that RAG transposition can be suppressed by the "non-core" C-terminal portion of RAG2.[71-73] Interestingly, full-length RAG2 suppressed transposition of intact RSS substrates,[71-73] but had no effect on transposition of precleaved RSS substrates.[71,74] Since full-length RAG2 only suppressed transposition when coding DNA was present in the RAG1/2 complex, this finding suggested that the C-terminal portion of RAG2 blocks transposition of intact substrates by stably binding to coding ends within the cleaved signal complex, thereby occupying the target DNA binding site and preventing target capture (Fig. 3, Step 2).[71] While it is possible that the C-terminal portion of RAG2 also inhibits transposition at the step of target capture by the signal-end complex (Fig. 3, Step 3b),[72,73] two studies found that signal-end complexes containing full-length RAG2 were just as active in target capture[71] and transposition[71,74] as signal-end complexes containing core RAG2. Thus, we favor a model where full-length RAG2 inhibits transposition by stabilizing the cleaved signal complex and preventing subsequent target capture (Fig. 3, Step 2). However, even if the C-terminal portion of RAG2 does suppress transposition in this manner, coding ends are processed more rapidly than signal ends in vivo.[75] Therefore, since the signal-end complex, which is devoid of coding ends, must persist for some time in the cell, other layers of regulation must exist.

Regulation by GTP

One of these additional layers of regulation may be inhibition by GTP.[73] In vitro experiments revealed that at concentrations of 1 mM or higher, GTP (but not ATP, CTP, or UTP) inhibited RAG transposition by blocking target capture within the signal-end complex (Fig. 3, Step 3b). This inhibition was alleviated by introducing substitutions within a putative GTP-binding domain in RAG1. Several other transposases—such as the Tn7 transposable element,[76,77] bacteriophage Mu[78] and Drosophila P element transposase[79]—are similarly regulated by nucleotide-binding. However, since the average intracellular GTP concentration in cells is only 0.5 ± 0.2 mM[80] and GTP inhibits RAG transposition very weakly in this concentration range,[73] the extent to which this mechanism regulates transposition in vivo remains unclear.

Regulation by Disintegration

An additional layer of regulation may be the propensity of the RAG1/2 complex to resolve branched transposition intermediates via disintegration (Fig. 3, Steps 4b and 4d).[70] As mentioned earlier, the branched DNA molecules generated as a result of RAG transposition can be resolved in one of three ways: nick repair—leading to insertional mutagenesis (Fig. 4a); disintegration—regenerating both the blunt singal ends and the linear target DNA (Fig. 4b); or target DNA transesterification—generating hairpinned target ends and signal ends with 3′ overhangs of 5 nucleotides (Fig. 4c). However, at physiologic magnesium concentrations of 20-25 mM, disintegration seems to be favored over both target DNA transesterification and nick repair.[70] Thus, by essentially reversing the process of RAG transposition, RAG-catalyzed disintegration may very well contribute to the low levels of transposition observed in vivo.

Regulation by Target Site Selection

Target site selectivity by the RAG transposase might reduce the frequency of deleterious transposition events by channeling these insertions into relatively safe regions of the genome. Initial studies revealed that RAG transposition events are moderately biased towards GC-rich target sequences.[5,66] Subsequent studies confirmed this preference for GC-rich regions[50] and suggested that distorted DNA structures such as DNA mismatches,[50] hairpins[81,82] and single-strand—double-strand DNA junctions[81] can also act as preferred sites for RAG transposition. If RAG transposition events are targeted to these distorted DNA structures and if these structures are predominantly found within innocuous regions of the genome, then target site selectivity may help to limit the frequency of harmful transposition events. However, it remains unclear whether such distorted DNA structures are predominantly found within innocuous regions of the genome. In addition, while this form of regulation may help to reduce deleterious transposition events, it would not limit the overall frequency of transposition in vivo.

Additional Potential Regulatory Mechanisms

Although the C-terminus of RAG2, GTP, disintegration and target site selectivity may help to suppress deleterious RAG transposition events in vivo, additional as-yet-undiscovered regulatory mechanisms must also exist. That is, at physiological concentrations of 20-25 mM Mg^{2+}, 5 μM Ca^{2+} and 0.5 mM GTP, the C-terminus of RAG2 inhibits transposition ~10-fold,[71,73] GTP inhibits transposition ~5-fold[73] and disintegration inhibits transposition ~10-fold.[70] Taken together, these regulatory mechanisms would suppress transposition ~500-fold. However, since RAG transposition occurs so robustly in vitro, this level of suppression is insufficient to explain the low frequency of transposition observed in vivo.[3,4,67] Here, we will speculate about additional potential regulatory mechanisms for suppressing RAG transposition in vivo.

Coding DNA May Assist in Reducing the Frequency of Interchromosomal Transposition

The requirement for both V(D)J cleavage and coding end release prior to target capture suggests that coding DNA can inhibit RAG transposition. As discussed above, one way that coding DNA can suppress transposition is by occupying the non-RSS DNA binding site of the RAG1/2 complex, thereby preventing the RAG1/2 complex from binding target DNA and committing to the transposition pathway. However, coding ends could also help to prevent deleterious transposition events by temporarily tethering signal ends to the antigen receptor loci, thereby reducing the length of time that the signal end complex has to freely diffuse through the cell and capture interchromosomal target DNA. That is, since it appears that chromosomes each occupy their own distinct territories within the nucleus,[83,84] newly generated signal ends would initially be positioned away from other chromosomes. Consequently, the RAG1/2 complex bound to these signal ends would only be able to bind intrachromosomal target DNA. Given enough time, the signal end complex could randomly diffuse through the nucleus and potentially encounter interchromosomal target DNA. Yet, since signal ends are repaired to form signal joints at the G1/S transition,[75] there is a finite window of opportunity for the signal end complex to diffuse and capture interchromosomal target DNA. Moreover, since the RAG1/2 complex holds on to postcleavage coding ends within the cleaved signal complex, coding end-binding tethers the RAG transposase to the originating antigen receptor locus for a period of time, thereby reducing the length of time that the RAG transposase has to encounter interchromosomal target DNA. In this way, coding end-binding can help to minimize potentially harmful interchromosomal translocations by biasing RAG transposition towards intrachromosomal targets.

Intriguingly, it has been shown that whereas XRCC4$^{-/-}$p53$^{-/-}$ mice (which are generally deficient in NHEJ) develop progenitor B-cell lymphomas harboring interchromosomal IgH:c-myc translocations,[85] Artemis$^{-/-}$p53$^{-/-}$ mice (which are specifically deficient in coding end repair) develop progenitor B-cell lymphomas harboring intrachromosomal IgH:N-myc translocations.[86] Although it hasn't been determined whether these translocations are derived from RAG transposition events, this finding suggests that hand-off of coding ends from the RAG1/2 cleaved signal complex to Artemis could be an important step in determining whether the RAG proteins undergo intrachromosomal or interchromosomal transposition. Perhaps, in the absence of normal coding end repair by Artemis, the RAG1/2 cleaved signal complex persists in the cell, thereby giving the RAG1/2 transposase a much shorter window of opportunity to capture interchromosomal target DNA. If this short window of opportunity is not long enough for the RAG transposase to diffuse through the nucleus and come into contact with other chromosomes (e.g., the c-myc locus), then the only target DNA that would be readily available to the RAG transposase would be intrachromosomal (e.g., the N-myc locus). Thus, in the absence of Artemis, persistent coding end-binding within the RAG1/2 cleaved signal complex could bias RAG transposition events towards intrachromosomal targets. But even in the presence of Artemis, coding end-binding may generally reduce the length of time that the RAG transposase has to encounter interchromosomal target DNA, thereby helping to minimize potentially harmful interchromosomal RAG transposition events.

Mammalian NHEJ Proteins May Inhibit RAG Transposition

Although coding DNA may generally block RAG transposition by preventing target capture and the tethering effect of coding ends may aid in specifically preventing interchromosomal RAG transposition, coding ends are repaired more rapidly than signal ends in vivo,[75] suggesting that the inhibitory effect of coding ends is likely to be transient. However, the requirement for coding end release prior to target capture raises the possibility of sustained inhibition of RAG transposition by the NHEJ proteins. That is, since the RAG1/2 cleaved signal complex must interact with the NHEJ proteins during the hand-off of coding ends, NHEJ proteins have an opportunity to influence the decision of the RAG1/2 complex to channel signal ends towards signal joint formation (Fig. 3, Step 3a) or target capture (Fig. 3, Step 3b) and subsequent transposition. While interacting with the RAG1/2

cleaved signal complex, perhaps the NHEJ proteins induce a conformational change in the RAG proteins (represented in Fig. 3 as a color change from white to gray at Step 2), that favors signal joint formation. This conformational change could either close the target DNA binding pocket or simply induce a conformation that favors subsequent interaction between the signal end complex and the NHEJ proteins. In this way, NHEJ proteins could help to inhibit transposition in vivo. Along these lines, it is interesting to note that when RAG1 and RAG2 are expressed in *Saccharomyces cerevisiae*, transposition occurs at least as frequently as signal joint formation.[68] Since yeast lack Artemis and DNA-PK_{CS} homologues, these two proteins may suppress RAG transposition in mammalian cells. Furthermore, although yeast do have homologues of mammalian Ku70,[87] Ku80,[88,89] XRCC4,[90] XLF[91,92] and DNA Ligase IV,[93] these factors are rather poorly conserved, with <25% identity between yeast and humans.[87,89-93] Given the low degree of conservation between yeast and human NHEJ factors, it seems plausible that mammalian Ku70, Ku80, XRCC4, XLF and DNA Ligase IV may suppress RAG transposition even though their yeast homologues do not.

Methylated Histone-Binding May Inhibit RAG Transposition

Given that the C-terminal portion of RAG2 has been demonstrated to suppress transposition in vitro,[71-73] it seems reasonable to hypothesize that other RAG-intrinsic regulatory mechanisms may exist to inhibit transposition in vivo. Recently, it has been demonstrated that a plant homeodomain (PHD) finger present in the C-terminal portion of RAG2 recognizes histone H3 when it is either trimethylated on lysine 4 (H3K4me3)[29,94] or when it is simultaneously symmetrically dimethylated on arginine 2 and trimethylated on lysine 4 (H3R2me2s/K4me3).[95] Furthermore, it has been shown that reducing either the levels of H3K4 methylation or the ability of RAG2 to bind H3K4me3 impairs V(D)J recombination, indicating that recognition of methylated histone H3 is important for V(D)J recombination in vivo.[29] Since V(D)J recombination is regulated by methylated histone-binding, RAG transposition may also be regulated by methylated H3-binding. Methylated H3-binding may allosterically inhibit the transposition activity of the RAG transposase by either closing the target DNA binding pocket and blocking target capture (Step 3b) or by inhibiting strand transfer (Steps 4a/4c). Alternatively, methylated H3-binding may regulate target site selection and direct RAG transposition into regions of the genome that are enriched for H3K4me3 and/or H3R2me2s/K4me3. Interestingly, many retrotransposons contain chromodomains[96-98]—a protein module that mediates interactions with methylated histone proteins—suggesting that regulation by methylated histone-binding may be a general feature of many different transposons. In the future, it will be interesting to see whether RAG transposition is, indeed, regulated by recognition of H3K4me3 and/or H3R2me2s/K4me3.

Regulation by Other Trans-Acting Factors

Finally, it worth noting that in much the same way that the NHEJ proteins may act to suppress RAG transposition, other, as-yet-unidentified, trans-acting factors may also regulate RAG transposition in vivo. These factors may stably interact with the RAG transposase to directly inhibit one of the steps in the transposition pathway. Alternatively, they may transiently interact with the RAG transposase to modify one (or both) of the RAG proteins in such a way as to inhibit transposition. These modifications could be either covalent (e.g., protein phosphorylation) or noncovalent (e.g., protein remodeling by an ATP-dependent molecular chaperone). In any case, it seems likely that additional regulatory factors may be involved in the inhibition of RAG transposition.

Conclusion

As described above, elucidation of the RAG transposition pathway has provided a useful conceptual framework within which to understand the regulation of RAG transposition. Although several regulatory mechanisms have already been identified, it seems likely that new forms of regulation will come to light in the next few years. In the future, it will be interesting to test how these various regulatory mechanisms interact with each other to suppress RAG transposition in vivo.

References

1. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell 2002; 109 Suppl:S45-55.
2. Gellert M. V(D)J recombination: RAG proteins, repair factors and regulation. Annu Rev Biochem 2002; 71:101-132.
3. Messier TL, O'Neill JP, Hou SM et al. In vivo transposition mediated by V(D)J recombinase in human T-lymphocytes. EMBO J 2003; 22(6):1381-1388.
4. Reddy YV, Perkins EJ, Ramsden DA. Genomic instability due to V(D)J recombination-associated transposition. Genes Dev 2006; 20(12):1575-1582.
5. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 1998; 94(4):463-470.
6. Tonegawa S. Somatic generation of antibody diversity. Nature 1983; 302(5909):575-581.
7. Akamatsu Y, Tsurushita N, Nagawa F et al. Essential residues in V(D)J recombination signals. J Immunol 1994; 153(10):4520-4529.
8. Akira S, Okazaki K, Sakano H. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. Science 1987; 238(4830):1134-1138.
9. Hesse JE, Lieber MR, Mizuuchi K et al. V(D)J recombination: a functional definition of the joining signals. Genes Dev 1989; 3(7):1053-1061.
10. Montalbano A, Ogwaro KM, Tang A et al. V(D)J recombination frequencies can be profoundly affected by changes in the spacer sequence. J Immunol 2003; 171(10):5296-5304.
11. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248(4962):1517-1523.
12. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell 1989; 59(6):1035-1048.
13. Mombaerts P, Iacomini J, Johnson RS et al. RAG-1-deficient mice have no mature B- and T-lymphocytes. Cell 1992; 68(5):869-877.
14. Oettinger MA. Activation of V(D)J recombination by RAG1 and RAG2. Trends Genet 1992; 8(12):413-416.
15. Shinkai Y, Rathbun G, Lam KP et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell 1992; 68(5):855-867.
16. McBlane JF, van Gent DC, Ramsden DA et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 1995; 83(3):387-395.
17. Sadofsky MJ, Hesse JE, McBlane JF et al. Expression and V(D)J recombination activity of mutated RAG-1 proteins. Nucleic Acids Res 1993; 21(24):5644-5650.
18. Kirch SA, Sudarsanam P, Oettinger MA. Regions of RAG1 protein critical for V(D)J recombination. Eur J Immunol 1996; 26(4):886-891.
19. Silver DP, Spanopoulou E, Mulligan RC et al. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. Proc Natl Acad Sci USA 1993; 90(13):6100-6104.
20. Cuomo CA, Oettinger MA. Analysis of regions of RAG-2 important for V(D)J recombination. Nucleic Acids Res 1994; 22(10):1810-1814.
21. Sadofsky MJ, Hesse JE, Gellert M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. Nucleic Acids Res 1994; 22(10):1805-1809.
22. Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu Rev Immunol 1999; 17:109-147.
23. Peixoto BR, Mikawa Y, Brenner S. Characterization of the recombinase activating gene-1 and 2 locus in the Japanese pufferfish, Fugu rubripes. Gene 2000; 246(1-2):275-283.
24. Akamatsu Y, Monroe R, Dudley DD et al. Deletion of the RAG2 C terminus leads to impaired lymphoid development in mice. Proc Natl Acad Sci USA 2003; 100(3):1209-1214.
25. Dudley DD, Sekiguchi J, Zhu C et al. Impaired V(D)J recombination and lymphocyte development in core RAG1-expressing mice. J Exp Med 2003; 198(9):1439-1450.
26. Elkin SK, Ivanov D, Ewalt M et al. A PHD finger motif in the C terminus of RAG2 modulates recombination activity. J Biol Chem 2005; 280(31):28701-28710.
27. Kirch SA, Rathbun GA, Oettinger MA. Dual role of RAG2 in V(D)J recombination: catalysis and regulation of ordered Ig gene assembly. EMBO J 1998; 17(16):4881-4886.
28. Liang HE, Hsu LY, Cado D et al. The "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B- and T-cell development. Immunity 2002; 17(5):639-651.
29. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450(7172):1106-1110.
30. Gu Y, Seidl KJ, Rathbun GA et al. Growth retardation and leaky SCID phenotype of Ku70-deficient mice. Immunity 1997; 7(5):653-665.

31. Ouyang H, Nussenzweig A, Kurimasa A et al. Ku70 is required for DNA repair but not for T-cell antigen receptor gene recombination In vivo. J Exp Med 15 1997; 186(6):921-929.
32. Nussenzweig A, Chen C, da Costa Soares V et al. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. Nature 1996; 382(6591):551-555.
33. Zhu C, Bogue MA, Lim DS et al. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. Cell 1996; 86(3):379-389.
34. Gao Y, Chaudhuri J, Zhu C et al. A targeted DNA-PKcs-null mutation reveals DNA-PK-independent functions for KU in V(D)J recombination. Immunity 1998; 9(3):367-376.
35. Taccioli GE, Amatucci AG, Beamish HJ et al. Targeted disruption of the catalytic subunit of the DNA-PK gene in mice confers severe combined immunodeficiency and radiosensitivity. Immunity 1998; 9(3):355-366.
36. Moshous D, Callebaut I, de Chasseval R et al. Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. Cell 2001; 105(2):177-186.
37. Rooney S, Sekiguchi J, Zhu C et al. Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. Mol Cell 2002; 10(6):1379-1390.
38. Gao Y, Sun Y, Frank KM et al. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. Cell 1998; 95(7):891-902.
39. Frank KM, Sekiguchi JM, Seidl KJ et al. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. Nature 1998; 396(6707):173-177.
40. Grawunder U, Zimmer D, Fugmann S et al. DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes. Mol Cell 1998; 2(4):477-484.
41. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 2006; 124(2):301-313.
42. Buck D, Malivert L, de Chasseval R et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 2006; 124(2):287-299.
43. Zha S, Alt FW, Cheng HL et al. Defective DNA repair and increased genomic instability in Cernunnos-XLF-deficient murine ES cells. Proc Natl Acad Sci USA 2007; 104(11):4518-4523.
44. Dai Y, Kysela B, Hanakahi LA et al. Nonhomologous end joining and V(D)J recombination require an additional factor. Proc Natl Acad Sci USA 2003; 100(5):2462-2467.
45. Huye LE, Purugganan MM, Jiang MM et al. Mutational analysis of all conserved basic amino acids in RAG-1 reveals catalytic, step arrest and joining-deficient mutants in the V(D)J recombinase. Mol Cell Biol 2002; 22(10):3460-3473.
46. Lee GS, Neiditch MB, Salus SS et al. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. Cell 2004; 117(2):171-184.
47. Leu TM, Eastman QM, Schatz DG. Coding joint formation in a cell-free V(D)J recombination system. Immunity 1997; 7(2):303-314.
48. Qiu JX, Kale SB, Yarnell Schultz H et al. Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol Cell 2001; 7(1):77-87.
49. Ramsden DA, Paull TT, Gellert M. Cell-free V(D)J recombination. Nature 1997; 388(6641):488-491.
50. Tsai CL, Chatterji M, Schatz DG. DNA mismatches and GC-rich motifs target transposition by the RAG1/RAG2 transposase. Nucleic Acids Res 2003; 31(21):6180-6190.
51. Yarnell Schultz H, Landree MA, Qiu JX et al. Joining-deficient RAG1 mutants block V(D)J recombination in vivo and hairpin opening in vitro. Mol Cell 2001; 7(1):65-75.
52. Bredemeyer AL, Sharma GG, Huang CY et al. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature 2006; 442(7101):466-470.
53. Clatworthy AE, Valencia-Burton MA, Haber JE et al. The MRE11-RAD50-XRS2 complex, in addition to other nonhomologous end-joining factors, is required for V(D)J joining in yeast. J Biol Chem 2005; 280(21):20247-20252.
54. Jones JM, Gellert M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. EMBO J 2002; 21(15):4162-4171.
55. Mundy CL, Patenge N, Matthews AG et al. Assembly of the RAG1/RAG2 synaptic complex. Mol Cell Biol 2002; 22(1):69-77.
56. Curry JD, Geier JK, Schlissel MS. Single-strand recombination signal sequence nicks in vivo: evidence for a capture model of synapsis. Nat Immunol 2005; 6(12):1272-1279.
57. Hiom K, Gellert M. Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. Mol Cell 1998; 1(7):1011-1019.
58. van Gent DC, Mizuuchi K, Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration. Science 1996; 271(5255):1592-1594.

59. Ma Y, Pannicke U, Schwarz K et al. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell 2002; 108(6):781-794.
60. Goodarzi AA, Yu Y, Riballo E et al. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. EMBO J 2006; 25(16):3880-3889.
61. Grawunder U, Wilm M, Wu X et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature 1997; 388(6641):492-495.
62. Grawunder U, Zimmer D, Kulesza P et al. Requirement for an interaction of XRCC4 with DNA ligase IV for wild-type V(D)J recombination and DNA double-strand break repair in vivo. J Biol Chem 1998; 273(38):24708-24714.
63. Lu H, Pannicke U, Schwarz K et al. Length-dependent binding of human XLF to DNA and stimulation of XRCC4.DNA ligase IV activity. J Biol Chem 2007; 282(15):11155-11162.
64. Sakano H, Huppi K, Heinrich G et al. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 1979; 280(5720):288-294.
65. Thompson CB. New insights into V(D)J recombination and its role in the evolution of the immune system. Immunity 1995; 3(5):531-539.
66. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 1998; 394(6695):744-751.
67. Chatterji M, Tsai CL, Schatz DG. Mobilization of RAG-generated signal ends by transposition and insertion in vivo. Mol Cell Biol 2006; 26(4):1558-1568.
68. Clatworthy AE, Valencia MA, Haber JE et al. V(D)J recombination and RAG-mediated transposition in yeast. Mol Cell 2003; 12(2):489-499.
69. Matthews AG, Elkin SK, Oettinger MA. Ordered DNA release and target capture in RAG transposition. EMBO J 2004; 23(5):1198-1206.
70. Melek M, Gellert M. RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. Cell 2000; 101(6):625-633.
71. Elkin SK, Matthews AG, Oettinger MA. The C-terminal portion of RAG2 protects against transposition in vitro. EMBO J 2003; 22(8):1931-1938.
72. Swanson PC, Volkmer D, Wang L. Full-length RAG-2 and not full-length RAG-1, specifically suppresses RAG-mediated transposition but not hybrid joint formation or disintegration. J Biol Chem 2004; 279(6):4034-4044.
73. Tsai CL, Schatz DG. Regulation of RAG1/RAG2-mediated transposition by GTP and the C-terminal region of RAG2. EMBO J 2003; 22(8):1922-1930.
74. Jiang H, Ross AE, Desiderio S. Cell cycle-dependent accumulation in vivo of transposition-competent complexes between recombination signal ends and full-length RAG proteins. J Biol Chem 2004; 279(9):8478-8486.
75. Ramsden DA, Gellert M. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. Genes Dev 1995; 9(19):2409-2420.
76. Peters JE, Craig NL. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. Genes Dev 2001; 15(6):737-747.
77. Stellwagen AE, Craig NL. Mobile DNA elements: controlling transposition with ATP-dependent molecular switches. Trends Biochem Sci 1998; 23(12):486-490.
78. Yamauchi M, Baker TA. An ATP-ADP switch in MuB controls progression of the Mu transposition pathway. EMBO J 1998; 17(18):5509-5518.
79. Kaufman PD, Rio DC. P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell 1992; 69(1):27-39.
80. Traut TW. Physiological concentrations of purines and pyrimidines. Mol Cell Biochem 1994; 140(1):1-22.
81. Lee GS, Neiditch MB, Sinden RR et al. Targeted transposition by the V(D)J recombinase. Mol Cell Biol 2002; 22(7):2068-2077.
82. Posey JE, Pytlos MJ, Sinden RR et al. Target DNA structure plays a critical role in RAG transposition. PLoS Biol 2006; 4(11):e350.
83. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet 2001; 2(4):292-301.
84. Parada L, Misteli T. Chromosome positioning in the interphase nucleus. Trends Cell Biol 2002; 12(9):425-432.
85. Gao Y, Ferguson DO, Xie W et al. Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. Nature 2000; 404(6780):897-900.
86. Rooney S, Sekiguchi J, Whitlow S et al. Artemis and p53 cooperate to suppress oncogenic N-myc amplification in progenitor B-cells. Proc Natl Acad Sci USA 2004; 101(8):2410-2415.

87. Feldmann H, Winnacker EL. A putative homologue of the human autoantigen Ku from Saccharomyces cerevisiae. J Biol Chem 1993; 268(17):12895-12900.
88. Boulton SJ, Jackson SP. Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance. Nucleic Acids Res 1996; 24(23):4639-4648.
89. Milne GT, Jin S, Shannon KB et al. Mutations in two Ku homologs define a DNA end-joining repair pathway in Saccharomyces cerevisiae. Mol Cell Biol 1996; 16(8):4189-4198.
90. Herrmann G, Lindahl T, Schar P. Saccharomyces cerevisiae LIF1: a function involved in DNA double-strand break repair related to mammalian XRCC4. EMBO J 1998; 17(14):4188-4198.
91. Callebaut I, Malivert L, Fischer A et al. Cernunnos interacts with the XRCC4 x DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. J Biol Chem 2006; 281(20):13857-13860.
92. Hentges P, Ahnesorg P, Pitcher RS et al. Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos. J Biol Chem 2006; 281(49):37517-37526.
93. Schar P, Herrmann G, Daly G et al. A newly identified DNA ligase of Saccharomyces cerevisiae involved in RAD52-independent repair of DNA double-strand breaks. Genes Dev 1997; 11(15):1912-1924.
94. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27(4):561-571.
95. Ramon-Maiques S, Kuo AJ, Carney D et al. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. Proc Natl Acad Sci USA 2007; 104(48):18993-18998.
96. Hizi A, Levin HL. The integrase of the long terminal repeat-retrotransposon tf1 has a chromodomain that modulates integrase activities. J Biol Chem 2005; 280(47):39086-39094.
97. Malik HS, Eickbush TH. Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. J Virol 1999; 73(6):5186-5190.
98. Nagaki K, Neumann P, Zhang D et al. Structure, divergence and distribution of the CRR centromeric retrotransposon family in rice. Mol Biol Evol 2005; 22(4):845-855.

Chapter 3

Recent Insights into the Formation of RAG-Induced Chromosomal Translocations

Vicky L. Brandt and David B. Roth*

Abstract

Chromosomal translocations are found in many types of tumors, where they may be either a cause or a result of malignant transformation. In lymphoid neoplasms, however, it is clear that pathogenesis is initiated by any of a number of recurrent DNA rearrangements. These particular translocations typically place an oncogene under the regulatory control of an Ig or TCR gene promoter, dysregulating cell growth, differentiation, or apoptosis. Given that physiological DNA rearrangements (V(D)J and class switch recombination) are integral to lymphocyte development, it is critical to understand how genomic stability is maintained during these processes. Recent advances in our understanding of DNA damage signaling and repair have provided clues to the kinds of mechanisms that lead to V(D)J-mediated translocations. In turn, investigations into the regulation of V(D)J joining have illuminated a formerly obscure pathway of DNA repair known as alternative NHEJ, which is error-prone and frequently involved in translocations. In this chapter we consider recent advances in our understanding of the functions of the RAG proteins, RAG interactions with DNA repair pathways, damage signaling and chromosome biology, all of which shed light on how mistakes at different stages of V(D)J recombination might lead to leukemias and lymphomas.

Introduction

Lymphoid neoplasms are among the most common malignancies in humans; mysteriously, they have become increasingly common in both adults and children over the past two decades, with the incidence of non-Hodgkin's lymphoma alone having doubled.[1] A number of factors are implicated in the etiology of these disorders, including ionizing radiation, chemical exposures, viral infection, autoimmune disease and acquired immunodeficiencies. Some of these conditions might directly create genetic mutations that initiate tumorigenesis; others may simply promote a favorable immune milieu by chronic antigenic stimulation or immunosuppression. It is fairly certain, however, that many lymphoid neoplasms are born of chromosomal translocations involving antigen receptor loci.[2,3] Up to 90% of cases of non-Hodgkin's lymphoma, for instance, bear such translocations.[1] These aberrant rearrangements most often exert their oncogenic effects by placing an oncogene under the regulatory control of a highly expressing Ig or TCR gene promoter, thereby dysregulating cell differentiation, proliferation, or survival.[3-5] Translocations also commonly

*Corresponding Author: David B. Roth—Department of Pathology and Program in Molecular Pathogenesis, The Helen L. and Martin S. Kimmel Center for Biology and Medicine at the Skirball Institute for Biomolecular Medicine, New York University School of Medicine, New York, NY 10016, USA. Email: roth@saturn.med.nyu.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

fuse the coding sequences of two different genes, which then encode chimeric oncoproteins that activate oncogenic transcriptional programs.[6] Both types of events frequently bear signs of having originated through some error in V(D)J recombination, the process by which antigen receptor genes are rearranged.[2,3,7,8]

V(D)J recombination can be thought of as a special case of targeted, strictly regulated genomic instability. There are seven antigen receptor loci that encode the T-cell receptor (TCR) α, β, γ and δ chains and the immunoglobulin (Ig) H and L (κ and λ) chains. Groups of V, D and J coding segments are arrayed along the loci, flanked by recombination signal sequences (RSS). The lymphoid-specific recombinase, consisting of RAG1 and RAG2 (the protein products of the recombination activating genes 1 and 2), selects a pair of signal sequences that may be many kilobases apart, cleaves the DNA at the signal sequence borders, and the resulting DNA double-strand breaks are joined by the ubiquitous nonhomologous end joining (NHEJ) proteins. Since antigen receptor gene rearrangement entails breaking and rejoining the chromosome several times before a complete Ig or TCR molecule can be expressed on the cell surface, the creation of a diverse repertoire of antigen receptors violates genomic integrity as a matter of course. It has been estimated that, each day, the human body creates 1×10^{11} B-cells.[9] Granted, most of these newly generated cells die because they form nonfunctional or self-reactive antigen receptors. Even so, an estimated 9×10^9 cells survive this process every day.[9] These numbers are staggeringly large. An error rate of less than a thousandth of a percent would still yield a large number of cells bearing potentially oncogenic translocations. How is it that leukemias and lymphomas do not overcome us all? The mechanisms that preserve genomic integrity during rearrangement must be unusually reliable, multiply redundant, or both.

In fact, the obvious risks attendant upon sequential cutting and pasting of gene fragments are mitigated by numerous restrictions on the process, many of which have only just been appreciated (and many others of which, no doubt, remain to be discovered). Regulation of recombination requires deft orchestration of chromatin changes, trans-acting factors, transcription, selection of substrates for DNA cleavage and DNA double-strand break (DSB) repair machinery. There are excellent reviews in this volume that do greater justice to the topic of accessibility than we could in this chapter (see also refs. 10-12). Our focus will be on recent work elucidating the molecular mechanisms for maintaining the fidelity of DSB repair. We will begin the chapter by outlining the salient features of the V(D)J reaction. We will then consider those stages where mistakes often occur, with a focus on mechanisms that can lead, in theory at least, to translocations.

Overview of the V(D)J Recombination Reaction

Key steps in the reaction are outlined below. For comprehensive and elegant descriptions of the biochemistry, see references 7, 13 and 14.

The recombination signal sequences (RSS) that flank the V, D and J segments consist of conserved heptamer and nonamer elements separated by an intervening spacer of either 12 or 23 nucleotides. These recognition sequences are referred to as 12-RSS or 23-RSS, and efficient recombination requires that two complementary RSS (a 12/23 pair) be synapsed before cleavage can proceed.[15-17] The heptamer has the palindromic consensus sequence CACAGTG, but variations are common and the extent of deviation from the consensus influences the efficiency with which a site is cleaved. The AT-rich nonamer sequence is less conserved but still important for recombination[18], and even the spacer sequences influence the selection of an RSS.[19-22]

The RSS are recognized by the lymphoid-specific proteins RAG1 and RAG2 ("recombination activating genes 1 and 2"[23]), which together form a complex we will refer to as the V(D)J or RAG recombinase. HMGB1 (high mobility group box 1), a nonspecific DNA bending protein, facilitates synaptic complex formation and cleavage.[24,25] The RAG proteins nick one DNA strand precisely between the RSS heptamer and the coding segment. This generates a free 3'OH that is used to attack the opposite strand in a transesterification reaction, forming a double-strand break (DSB). The result is that the synapsed pair of RSS/coding segments yields four free DNA ends: two covalently sealed (hairpin) coding ends and two signal ends that terminate in a flush double-strand break.[26-30]

After coupled cleavage, the RAG proteins hold the DNA ends in a postcleavage complex, aligning them for proper joining by the nonhomologous end joining (NHEJ) machinery. The blunt-ended RSS undergo direct ligation (generally with no base loss) to form a signal joint, which is usually deleted as an extrachromosomal circular product that is lost during cell division. Less frequently, the orientation of the coding segments necessitates inversional recombination, in which the signal joint is retained in the chromosome. There is no known immunological function for signal joints, but in cases of inversional recombination their formation is necessary for preserving genomic integrity. Ligation of the two coding ends produces a coding joint that encodes the variable portion of the antigen receptor protein. Coding joints are typically imprecise, as the coding end hairpins must first be opened and often undergo loss or addition of nucleotides during processing. This junctional variability contributes further to antigen receptor diversity and is considered characteristic of repair by nonhomologous end-joining.

Potential Mechanisms of RAG-Mediated Translocations

Errors in recombination can be broadly classified into two categories. Those occurring during the early stage of the reaction (site selection and cleavage) can be conceptualized as cases of mistaken identity: they involve either (1) mixing of authentic but inappropriate antigen receptor loci (e.g., TCRβ and TCRγ segments) in interlocus recombination, or (2) the misappropriation of sequences that fortuitously resemble RSS (cryptic RSS). One mechanism for preventing such errors involves regulation of substrate accessibility; we will discuss this and related regulatory controls relevant to each type of substrate selection error in the following section. Errors that take place in later stages of the reaction (joining) can instead be conceived as involving renegade double-strand breaks. Broken DNA ends created in the context of V(D)J recombination might escape normal DNA repair through defects in the RAG postcleavage complex, use of an inappropriate repair pathway, or an impaired DNA damage signaling response. Mechanisms that act to curtail aberrant repair will be considered in the context of these deficits in subsequent sections.

Mistaken Identities: Substrate Selection Errors

Interlocus Recombination

Normal V(D)J recombination is restricted by cell lineage (TCR loci rearrange in T-cells but not B-cells), developmental stage (e.g., TCRβ before TCRα) and, in many cells, to one allele (allelic exclusion). Since the RAG proteins, the RSS and the DNA repair machinery are the same in each case, this complex regulatory scheme depends in large part on the degree of accessibility allowed the recombinase to the various loci over time in different cells. For this reason, the packaging of TCR and Ig loci into chromatin differs in B- and T-cells and varies according to the activity of the loci, which is governed by developmental stage.

Nevertheless, some temporal overlap in the sequence of rearrangements does allow occasional interlocus (trans) recombination.[31-34] These rearrangements, which create a balanced translocation resulting in two derivative chromosomes, can generate functional chimeric receptor chains that appear in normal tissues.[33,34] As with recurrent oncogenic translocations, the system seems to favor rearrangements of particular sites: for example, it has been estimated that 1 in 10,000 normal human and mouse thymocytes carries the Dδ3-Jβ2.7 rearrangement.[32,35] These rearrangements, just like those that occur in *cis*, rely on RSS recognition, RAG-mediated cleavage and NHEJ repair. They are normal V(D)J reactions simply carried out with the wrong partner. Interlocus events do, however, exhibit recurrent base loss from signal joints[31,36] and difficulty forming coding joints.[37-39] These features suggest that trans rearrangements proceed through an abnormal pathway.

It is noteworthy that the incidence of interlocus recombination increases dramatically in cells bearing certain mutations (such as ATM deficiency) that predispose to lymphoid tumors.[32,40-42] These events have the appearance of simple substrate selection errors, but at least some of these rearrangements might arise from failures in DNA damage sensing and repair (see discussion of ATM defects below, in the section "The role of the DNA damage response in preventing translocations").

Cryptic RSS

The variability of RSS sequence entails considerable flexibility on the part of the RAG proteins. Unfortunately, this plasticity makes it possible for the RAG proteins to bind to fortuitous DNA sequences known as "cryptic RSS" that do not border antigen receptor gene segments but are sufficiently close to the consensus sequence to allow RAG recognition.[43,44] In one large review of oncogenic rearrangements from both B- and T-cell malignancies, most translocation breakpoints on the nonantigen receptor gene partner contained RSS-like sequences at or near the breakpoint, supporting "substrate selection error" as the responsible mechanism.[2] In addition, nontemplated nucleotides are frequently added to the junctions, suggesting TdT activity and therefore the involvement of V(D)J recombination.[2] The t(7; 9) (q34; q32) translocations found in T-cell lymphoblastic leukemia provide the clearest example. Chromosome 7 breakpoints are typically located at the RSS bordering Dβ segments, while breakpoints on chromosome 9 are flanked by consensus RSS heptamer sequences separated from AT-rich nonamer-like sequences by 11 or 12 base pairs.[45] The salient feature of substrate selection errors is that the V(D)J recombination reaction proceeds as normal except for partnering an RSS with an inauthentic sequence.

Preventing Errors by Controlling Accessibility

An RSS can deviate quite far from the consensus and still undergo recombination; Lewis et al defined the necessary features of cryptic RSS and suggested that even a weak signal, with a recombination frequency of 2×10^{-5} the canonical level, can have a physiological impact.[43] In light of estimates that the genome contains 10 million potential cryptic sites, approximately one every 1-2 Kb,[46] it is clear that RAG accessibility to target sites must be very tightly regulated.

In a prescient 1985 paper, Yancopoulos and Alt noted that rearranging segments are transcribed before (or coincident with) their activation for rearrangement and proposed that generating these germline transcripts altered chromatin structure so as to allow the recombinase access to a subset of appropriate substrates.[47] There are also other potential mechanisms for regulating locus accessibility that do not rely on transcription.[48] One approach to controlling access is through nucleosome packaging, which can block cleavage of specific RSS.[49] Proteins that enhance RAG interaction with RSSs[48,50,51] could conceivably recruit nucleosome remodeling complexes such as Swi/Snf that alter DNA-histone contacts within a nucleosome or alter the nucleosome's location.[52,53] The second approach is through covalent modifications of the tail domains of the histone proteins by acetylation of lysines, methylation of lysines and arginines, polyribosylation, serine phosphorylation and ubiquitylation.[54] Such posttranslational modifications can "open" chromatin by altering DNA-histone contacts within a nucleosome, histone-histone contacts between nucleosomes, or interactions between histones and other proteins. Accumulating evidence suggests that these reversible, epigenetic modifications comprise a "histone code" and that they associate with regulatory proteins known as code readers. Evolutionarily conserved domains within code-reader proteins bind to certain histone modifications with such specificity that they can distinguish the same modification at different residues (for example, trimethylation at K4 vs. K9).[54]

Several recent studies have shown that the plant homeodomain (PHD) finger, a methyl-lysine binding domain, serves as a code-reader: it can both promote and repress gene expression by interacting with trimethylated lysine 4 on histone 3 (H3K4).[55-58] Even more recently, the RAG2 PHD finger has been shown to recognize H3K4 trimethylation.[59-61] In these studies, the binding of RAG2 to H3K4 enhanced the selection and recombination of chromatinized gene segments in developing lymphocytes. The RAG complex, then, is not merely subject to chromatin structures determined by other factors, but must take an active role in recognizing substrates.

Other studies have shown that transcriptional cis-regulatory sequences, such as enhancers and promoters specific to each locus, are necessary for V(D)J recombination.[12,62] Furthermore, the RAG genes are regulated differently in B- and T-cells (for example, Foxp1 is required for B-cell-specific RAG expression[63]). Some DNA-binding transcription factors interact with RAG1/RAG2 and guide them to subsets of RSS; B-cell-specific V_H locus contraction, for instance, requires Pax5 to interact with both the V coding segments and the RAG complex.[64,65] The mechanisms of locus

contraction and looping remains poorly understood, but they are essential for promoting synapse formation between distal V and proximal D segments, which can be separated by distances of up to 3 megabases.[66] (In this regard, it is interesting to note that core RAG2 knock-in mice have difficulty with V to DJ rearrangements at the IgH and TCRβ loci.[67,68]) Whether nonantigen receptor loci are typically constrained by such complex regulatory schemes is not clear.

Signs That a Translocation Did Not Arise through Substrate Selection Error

Even granting the occasional chromatin loophole, three observations suggest that substrate selection errors do not account for the majority of RAG-mediated oncogenic translocations. First, many of the RSS-like sequences found at translocation breakpoints on the nonantigen receptor partner chromsome contain heptamers that are a poor match for the consensus, and a large fraction lack recognizable nonamer elements.[2,7] Previous work has shown that DNA cleavage in vivo requires both heptamer and nonamer; scrambling the nonamer or mutating a single critical nucleotide in the heptamer decreases cleavage by at least two orders of magnitude.[15,18,22,69] Therefore, the presence of sequences that deviate so much from the consensus on the partner (nonantigen receptor locus) chromosome might be merely coincidental.[2,3,7] The second argument against the use of some cryptic RSS in translocations is that the breakpoints are often not at the heptamer-coding flank border. This is incompatible with normal RAG-mediated cleavage, which is a very precise reaction. Finally, some translocations display short direct repeats,[8,70] suggesting that the cleavage event created a short single-stranded overhang. This, too, is inconsistent with normal cleavage by the V(D)J recombinase.

This is not to say that such events did not originate with a mistake in V(D)J recombination. If substrate selection error appears unlikely, there is an alternative model that better explains cases such as these. It is known as end donation and posits that the recombinase creates a double-strand break (DSB) at an authentic RSS that is then somehow joined to a random DSB that has been created through some unrelated process.[7] Until the past few years it has been difficult to conceive of a mechanism that would explain end donation, but recent work suggests that broken DNA ends created by RAG cleavage might escape their normal fate through defects in the RAG postcleavage complex, use of an inappropriate repair pathway, or an impaired DNA damage signaling response.

The Ends That Got Away: Errors in Joining

DSBs are potentially so damaging that cells have evolved complex networks of proteins to sense the presence and precise location of DNA damage, regulate the cell cycle and repair the breaks. Mounting evidence suggests that V(D)J recombination enjoys at least two layers of protection that even its DNA-rearranging cousin, class switch recombination, does not:[71] an end joining pathway that discourages translocations (classical NHEJ) and the RAG postcleavage complex, which is thought to ensure joining through this pathway and exclude other, error-prone repair. Yet another layer of protection is provided by ATM, part of the DNA damage signaling machinery, which may have a role in stabilizing the postcleavage complex but also can lead cells with unrepaired breaks to undertake apoptosis.

Genome Guardians: The Classical NHEJ Factors

The basic outline of NHEJ seems simple enough: a set of enzymes captures the two ends of the broken DNA molecule, a molecular bridge is formed to juxtapose the ends, and the break is religated.[72] In reality the process is rather complex and many aspects remain poorly understood (see refs. 72 and 73). A key component of NHEJ is the DNA-dependent protein kinase (DNA-PK) complex, which comprises the DNA-PK catalytic subunit (DNA-PKcs) and the Ku70 and Ku80 nuclear antigens.[74] Nonhomologous repair is initiated when the Ku70/80 heterodimer encircles a broken end,[75,76] creating a scaffold for the recruitment of other factors. Ku attracts DNA-PKcs to the break, where it might serve multiple roles, including the formation of a synaptic complex to bring the ends together.[72] Activated DNA-PKcs recruits XRCC4, DNA Ligase IV and Artemis. DNA-PKcs phosphorylation of Artemis converts the latter from an exonuclease to an endonuclease and allows it to open the hairpinned coding ends.[77,78] Since Artemis cannot process every type of

nonligatable end, other types of end-processing enzymes are also recruited. Polymerase activity, for example, is likely supplied by the DNA polymerase Mu, which associates with XRCC4, and terminal deoxynucleotidyl transferase (TdT) adds nontemplated nucleotides to increase junctional diversity.[79,80] Finally, XRCC4 and DNA Ligase IV ligate the ends.[81-83] The most recently discovered NHEJ factor, known as Cernunnos or XLF (for XRCC4-like factor), is also recruited by Ku and interacts with both XRCC4 and Ligase IV to ligate mismatched and noncohesive ends.[84-88] The order in which all these factors are recruited might be flexible, according to the specific nature of the break.[89]

Genetic ablation of Ku, DNA-PKcs, DNA Ligase IV, XRCC4, Artemis, or Cernunnos in mice prevents the completion of V(D)J recombination, arresting B- and T-cell development at an early stage and leading to a SCID (severe combined immunodeficiency) phenotype. The overall defect in DNA repair also produces sensitivity to ionizing radiation, a marked tendency to translocations and development of lymphoma (though in some cases, only on a p53-deficient background).[90-97] (By contrast, NHEJ-proficient mammalian cells reconstitute their chromosomes with remarkable accuracy after being exposed to doses of ionizing radiation large enough to induce massive chromosome fragmentation.[98,99]) Some NHEJ-deficient lines develop nonlymphoid tumors as well.[90,100,101] The discovery that a deficiency of NHEJ factors promotes oncogenesis revealed a crucial role for these proteins as genome guardians.[94,95]

Error-Prone End Joining: Alternative NHEJ

Despite their obvious defects in DNA repair, NHEJ-deficient mice (and humans[97,102,103]) can survive long enough to develop malignancy. The mouse tumors frequently show gene fusions between the IgH locus and c-Myc but can display many other nonreciprocal translocations. There must, then, be alternative mechanisms capable of repairing DSB without Ku, DNA-PKcs, Ligase IV, or XRCC4. And, in fact, there is, although it was not recognized as an alternative pathway when it was originally described in mammalian cells in the 1980s.[104-106]

At the time, it was known that eukaryotic cells are able to repair DNA ends by both homologous and nonhomologous means. In the case of V(D)J recombination intermediates, homology-based mechanisms seemed unlikely, as little or no homology is present between coding ends; moreover, rearranged coding segments underwent a curious addition and loss of nucleotides at the junction.[107] The mechanism for nonhomologous repair, however, had not yet been discovered and the field struggled to understand how "unrelated DNA ends are joined together willy-nilly with high efficiency".[104] The similarity of these junctions to coding joints hinted that the DNA breaks generated by the V(D)J recombinase might be repaired by the same mechanism.[106] Within several years, studies of V(D)J recombination in various radiosensitive cell lines made it possible to identify components of the NHEJ pathway.[108-112] Our understanding of NHEJ thus grew out of our understanding of V(D)J recombination—and because the wild-type RAG complex guides DNA ends to the classical pathway, not the alternative pathway (see below), the latter settled into quiet obscurity. Only recently, in fact, has it been realized that the two pathways are distinct.[113-115]

The hallmarks of junctions formed by alternative NHEJ are excessive deletions and a reliance on short sequence homologies (microhomologies).[106,113,115] Even blunt-ended plasmids in Ku80-deficient cells undergo resection and annealing of microhomologous sequences rather than simply being joined at the blunt ends.[115] It is worth noting that these microhomologies are present at oncogenic translocations from NHEJ-deficient cells.[96] Therefore, although alternative NHEJ provides enough repair activity to allow cell survival, it appears to be error-prone and predisposes the cell to genomic instability.

But if alternative NHEJ is relatively efficient, why does NHEJ deficiency virtually obliterate V(D)J recombination?

The RAG Postcleavage Complex Governs Choice of Repair Pathway

The observation that both nucleotide addition and deletion could occur prior to joining of coding ends indicated that the DNA ends must remain in one place long enough to allow processing by polymerases and endonucleases.[116] Thus, even before the discovery of RAG1 and RAG2, it

seemed that a stable protein-DNA complex must exist to allow the ends to be accessible to such modifying enzymes after cleavage.[116] When studies showed that cells deficient in Ku or DNA-PK could not resolve V(D)J intermediates, it seemed reasonable to think that, by analogy with the Mu transposase, a very stable postcleavage complex would make DNA ends inaccessible.[117] As the field's understanding of NHEJ repair grew, so did curiosity about how a RAG postcleavage complex might participate in joining.

Lacking a viable in vitro system to study joining, we turned to genetics. Separation-of-function mutants in RAG-1 and RAG-2 that are capable of cleavage but exhibit severe joining defects provided compelling evidence that the postcleavage complex serves a crucial function in joining both coding and signal ends.[118-120] These data lent support to the notion that the RAG proteins form a scaffold that holds the ends together to facilitate joining. Joining mutants could alter the architecture of the complex, facilitating premature release of ends or, conversely, creating a too-stable complex or hindering the recruitment of NHEJ factors.[118-121] Intriguingly, two RAG-1 mutants phenocopied NHEJ mutants: the rare joints they did manage to form exhibited the excessive deletions and short sequence homologies characteristic of alternative NHEJ.[118] These mutants led us to propose that the RAG proteins might function as genome guardians within the context of V(D)J recombination.

We pursued this hypothesis further by examining whether RAG-generated ends could be made available to repair pathways other than NHEJ. (Although homologous recombination and NHEJ predominate at different phases of the cell cycle, accumulating evidence suggests that they can act at the same time and even cooperate to repair a DSB.[73,122]) Using an in vivo system to assay for repair of signal ends by homologous recombination, Lee et al showed that two joining-impaired RAG1 mutants destabilize the RAG postcleavage complex, allowing the DNA ends to be available for repair by homologous recombination.[123] Wild-type postcleavage complexes, by contrast, stimulated no homologous recombination. This led us to propose a model in which the normally quite stable RAG postcleavage complex actively directs DNA ends to the NHEJ machinery for repair.[123] The question remained: how do the rare coding joints produced in NHEJ-deficient cells manage to be formed by the alternative NHEJ pathway?

Since the homologous recombination assay was unable to map the fate of coding ends and we had identified mutations in RAG2 that affected joining without destabilizing the postcleavage complex, we again took a genetic approach. We identified a truncated RAG2 allele that allows substantial coding and signal joint formation to occur in cells deficient for DNA-PKcs or XRCC4.[124] Junction sequences revealed a tendency toward large deletions and microhomology use. Surprisingly, this RAG2 mutant also revealed alternative NHEJ to be active even in wild-type cells.[124] These studies, along with work from the Alt and de Villartay labs studying the use of alternative NHEJ in class switch recombination,[125,126] make it clear that alternative NHEJ is quite robust, albeit error-prone. Thus, we have come full circle: V(D)J recombination allowed the discovery of classical NHEJ and now has brought attention back to alternative NHEJ.

Why is classical NHEJ less prone to translocations than the alternative pathway? Perhaps components of the classical NHEJ pathway interact with chromatin (or chromosome) components to maintain the chromosomal identity of broken ends (see below). In addition, studies of NHEJ have revealed that repair is biphasic: most repair occurs quite rapidly upon induction of a DSB, but there is a slow component that might correspond to alternative pathways and which continues at the same level when the classical pathway is disabled.[127] Thus, it seems the rapidity of classical NHEJ repair ensures that most DSBs are healed within a few hours; those lesions that cannot be repaired in this time will be subject to alternative end joining. It is conceivable that difficult-to-repair DSBs lingering in the nucleus might, over time, separate or drift to a different chromosome territory in the course of other cellular processes (but see below).

How Do Chromosome Ends Meet?

Mammalian chromosomes occupy discrete three-dimensional regions in the nucleus known as chromosome territories. These territories are not fixed, but are specific to different cell types.[128] In

order for a translocation to occur, there must be DSBs in (at least) two chromosomes at the same time; the DSBs must have escaped the normal repair mechanisms; the broken chromosome ends must physically meet and they must be illegitimately repaired. An obvious question arises: do the DSBs roam the nucleus, looking for a partner, or do they stay put?

Two hypotheses have been put forth. The breakage-first model posits that breaks are able to traverse the nuclear space, searching for potential partners, and come together to produce translocations. The contact-first model, on the other hand, proposes that since chromosomes occupy territories in the nucleus, breaks on distinct chromosomes will meet only if they occupy nearby or intermingling domains.[128] To test these possibilities, Soutoglou et al developed a cell system in which they could induce one DSB at a defined site and follow the fate of each of the damaged DNA ends in real time by observing specific fluorescent tags on either side of the break.[129] The authors demonstrated that a single DSB in mammalian cells is positionally stable, with only slight motion of the DNA break.[129] This stability required the end-binding Ku80/Ku70 heterodimer but, surprisingly, was independent of other DNA repair factors, the structural proteins H2AX and SMC1, the cohesin complex and even the Mre11 complex, which has been strongly implicated in anchoring ends. Whether other factors will turn out to be necessary for this immobilization of a break—or whether the cause of the breakage, or the number of breaks induced at the same time, influence this positional stability—remains to be seen.

These results have striking implications for understanding how translocations form in vivo. First, they demonstrate that chromosomal positional stability is related to genomic stability. (At least in mammals; yeast do not have chromosome territories. DSBs in yeast migrate to any of several small nuclear sites that act as damage repair centers.[130]) Second, the data support a contact-first model in mammalian cells and are consistent with the emerging motion that nonrandom, higher order spatial organization of chromosomes accounts in large part for the recurrence of specific translocations. Ten years ago, experiments showed that γ-irradiation of normal human lymphocytes induces translocations in chromosome pairs that have been observed in leukemias, suggesting that these chromosomes are near neighbors in lymphocytes.[131,132] Several frequent translocation partners, including Myc-Igh and BCR-ABL, have been found to exist in close spatial proximity to each other in normal cells before the formation of translocations.[128] The misjoining of proximally positioned chromosome regions supports the observed correlation between the degree of chromosome intermingling and the likelihood of translocations.[133] The frequency of translocations involving antigen receptor loci likely reflects the fact that more gene-rich chromosomes undergo less compaction and more intermingling.[133]

The Role of the DNA Damage Response in Preventing Translocations

The DNA damage sensing pathway was not initially thought to be involved in V(D)J recombination, as damage checkpoints are not activated during the process; in fact, it was assumed that the RAG postcleavage complex sequestered the DSB from the DNA damage sensing machinery. It thus came as a surprise to find that ATM, γ-H2AX and the Mre11 complex localize to RAG-mediated DNA breaks.[134,135] The mystery was deepened by the first studies to investigate whether these factors had any role in V(D)J recombination: the answer, apparently, was no.[136,137] Further probing unearthed a greater tendency to TCR α/δ interlocus recombination in mice deficient for ATM, Mre11, Nbs1, or 53BP1.[42,138-141] Mice deficient in ATM, Rad50, or H2AX develop thymic lymphomas, as do H2AX- and Mre11-deficient mice on a p53 null background.[136-139] Many of these tumors harbor translocations thought to derive from errors in V(D)J recombination, and tumorigenesis is reduced or delayed in mice when ATM deficiency is coupled with RAG1 or RAG2 deficiency.[142,143] Mutations in ATM, Nbs1 and Mre11 cause Ataxia-Telangiectasia, Nijmegen Breakage syndrome and Ataxia-Telangiectasia-Like disorder, respectively; like the mice, patients with these diseases have a predisposition to lymphoid malignancies and harbor frequent translocations between the TCR and Ig loci.

Recent studies provide insight into the role played by ATM (and perhaps, by extension, other damage sensors) in V(D)J recombination and why this role is virtually invisible under normal

circumstances. In addition to its newly discovered role in stabilizing DSB complexes during V(D)J recombination,[144] ATM has a checkpoint function to prevent the propagation of DSBs caused either by RAG or low-dose gamma irradiation to daughter cells.[145] Callen and colleagues posit that $ATM^{-/-}$ lymphocytes that fail primary V(D)J assembly, leaving a DSB on one allele, can still achieve productive rearrangement through independent recombination of the second allele. The presence of the DSB in ATM-deficient cells would not prevent pre-B-cells from undergoing the maturational process. Therefore, DSBs produced in precursor cells would persist in mature B-cells in peripheral lymphoid tissues, where they would then undergo class switching and be subject to further (AID-mediated) DNA breakage.[145] The initial RAG-mediated break could persist for many days, ultimately to be joined to another chromosome in a progeny cell.

This model puts an interesting twist on extant models of how chromosome ends meet in the nucleus and undergo misrepair, forming a translocation. The work of Callen and colleagues supports a contact-first model but suggests that a DSB could migrate from its original position in the chromosome territories and participate in a repair event with another chromosome broken in a progeny cell.[145] One might think of this as diachronic end donation. With regard to physiological relevance, it is striking that up to 50% of mantle cell lymphomas have mutations or deletions in ATM.[146] Callen et al suggest that ATM mutation is likely to be an early event in the malignant transformation.[145]

The foregoing studies emphasize that creating (or preventing) a translocation is a complex process. One has to consider not only the nature of repair factors and the ordered assembly and disassembly of DNA-protein complexes, but the fact that these processes take place in three dimensions and over time. Understanding the spatiotemporal regulation of these repair processes and their coordination with chromosome dynamics, changes in chromatin structure, DNA damage signaling, the cell cycle and other physiological processes represents one of the major challenges to unraveling the puzzle of aberrant V(D)J recombination events. Indeed, the recent discovery that over 700 proteins interact with ATM and ATR in the DNA damage response[147] indicates that this story is likely to get much more complicated.

References

1. Fisher SG, Fisher RI. The epidemiology of nonHodgkin's lymphoma. Oncogene 2004; 23(38):6524-6534.
2. Tycko B, Sklar J. Chromosomal translocations in lymphoid neoplasia: a reappraisal of the recombinase model. Cancer Cells 1990; 2:1-8.
3. Vanasse GJ, Concannon P, Willerford DM. Regulated genomic instability and neoplasia in the lymphoid lineage. Blood 1999; 94(12):3997-4010.
4. Kirsch IR, Morton CC, Nakahara K et al. Human immunoglobulin heavy chain genes map to a region of translocations in malignant B lymphocytes. Science 1982; 216(4543):301-303.
5. Dalla-Favera R, Bregni M, Erikson J et al. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci USA 1982; 79(24):7824-7827.
6. Look AT. Oncogenic transcription factors in the human acute leukemias. Science 1997; 278(5340):1059-1064.
7. Lewis SM. The mechanism of V(D)J joining: Lessons from molecular, immunological and comparative analyses. Adv Immunol 1994; 56:27-150.
8. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B-cell lymphomas. Oncogene 2001; 20(40):5580-5594.
9. Saada R, Weinberger M, Shahaf G et al. Models for antigen receptor gene rearrangement: CDR3 length. Immunol Cell Biol 2007; 85(4):323-332.
10. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
11. Schlissel MS. Regulating antigen-receptor gene assembly. Nat Rev Immunol 2003; 3(11):890-899.
12. Jung D, Giallourakis C, Mostoslavsky R et al. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annu Rev Immunol 2006; 24:541-570.
13. Gellert M. V(D)J recombination: RAG proteins, repair factors and regulation. Annu Rev Biochem 2002; 71:101-132.

14. Fugmann SD, Lee AI, Shockett PE et al. The RAG proteins and V(D)J recombination: complexes, ends and transposition. Annu Rev Immunol 2000; 18:495-527.
15. Steen SB, Gomelsky L, Roth DB. The 12/23 rule is enforced at the cleavage step of V(D)J recombination in vivo. Genes to Cells 1996; 1(6):543-553.
16. Eastman QM, Leu TMJ, Schatz DG. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. Nature 1996; 380:85-88.
17. van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. Cell 1996; 85:107-113.
18. Hesse JE, Lieber MR, Mizuuchi K et al. V(D)J recombination: a functional definition of the joining signals. Genes Dev 1989; 3:1053-1061.
19. Nadel B, Tang A, Escuro G et al. Sequence of the spacer in the recombination signal sequence affects V(D)J rearrangement frequency and correlates with nonrandom Vkappa usage in vivo. J Exp Med 1998; 187(9):1495-1503.
20. Bassing CH, Alt FW, Hughes MM et al. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. Nature 2000; 405(6786):583-586.
21. Feeney AJ, Goebel P, Espinoza CR. Many levels of control of V gene rearrangement frequency. Immunol Rev 2004; 200:44-56.
22. Swanson PC. The bounty of RAGs: recombination signal complexes and reaction outcomes. Immunol Rev 2004; 200:90-114.
23. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248:1517-1523.
24. van Gent DC, Hiom K, Paull TT et al. Stimulation of V(D)J cleavage by high mobility group proteins. EMBO J 1997; 16(10):2665-2670.
25. Dai Y, Wong B, Yen YM et al. Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination. Mol Cell Biol 2005; 25(11):4413-4425.
26. Roth DB, Nakajima PB, Menetski JP et al. V(D)J recombination in mouse thymocytes: Double-strand breaks near T-cell receptor delta rearrangement signals. Cell 1992; 69:41-53.
27. Roth DB, Menetski JP, Nakajima PB et al. V(D)J recombination: Broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell 1992; 70:983-991.
28. Roth DB, Zhu C, Gellert M. Characterization of broken DNA molecules associated with V(D)J recombination. Proc Natl Acad Sci USA 1993; 90:10788-10792.
29. Schlissel M, Constantinescu A, Morrow T et al. Double-strand signal sequence breaks in V(D)J recombination are blunt, 5'-phosphorylated, RAG-dependent and cell cycle regulated. Genes Dev 1993; 7:2520-2532.
30. McBlane JF, van Gent DC, Ramsden DA et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. Cell 1995; 83:387-395.
31. Tycko B, Palmer JD, Sklar J. T-cell receptor gene trans-rearrangements: chimeric gamma delta genes in normal lymphoid tissues. Science 1989; 245:1242-1246.
32. Kobayashi Y, Tycko B, Soreng AL et al. Transrearrangements between antigen receptor genes in normal human lymphoid tissues and in ataxia telangiectasia. J Immunol 1991; 147:3201-3209.
33. Davodeau F, Peyrat MA, Gaschet J et al. Surface expression of functional T-cell receptor chains formed by interlocus recombination on human T-lymphocytes. J Exp Med 1994; 180(5):1685-1691.
34. Bailey SN, Rosenberg N. Assessing the pathogenic potential of the V(D)J recombinase by interlocus immunoglobulin light-chain gene rearrangement. Mol Cell Biol 1997; 17(2):887-894.
35. Marculescu R, Le T, Simon P et al. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. J Exp Med 2002; 195(1):85-98.
36. Garcia IS, Kaneko Y, Gonzalez-Sarmiento R et al. A study of chromosome 11p13 translocations involving TCR beta and TCR delta in human T-cell leukaemia. Oncogene 1991; 6(4):577-582.
37. Han J-O, Steen SB, Roth DB. Intermolecular V(D)J recombination is prohibited specifically at the joining step. Mol Cell 1999; 3:331-338.
38. Tevelev A, Schatz DG. Intermolecular V(D)J recombination. J Biol Chem 2000; 275(12):8341-8348.
39. Agard EA, Lewis SM. Postcleavage sequence specificity in V(D)J recombination. Mol Cell Biol 2000; 20(14):5032-5040.
40. Lipkowitz S, Stern MH, Kirsch IR. Hybrid T-cell receptor genes formed by interlocus recombination in normal and ataxia-telangiectasia lymphocytes. J Exp Med 1990; 172(2):409-418.
41. Kirsch IR, Lipkowitz S. A measure of genomic instability and its relevance to lymphomagenesis. Cancer Res 1992; 52:5545s-5546s.
42. Theunissen JW, Kaplan MI, Hunt PA et al. Checkpoint failure and chromosomal instability without lymphomagenesis in Mre11(ATLD1/ATLD1) mice. Mol Cell 2003; 12(6):1511-1523.

43. Lewis SM, Agard E, Suh S et al. Cryptic signals and the fidelity of V(D)J joining. Mol Cell Biol 1997; 17(6):3125-3136.
44. Zhang M, Swanson PC. V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies. J Biol Chem 2008:283(11):6717-27.
45. Tycko B, Reynolds TC, Smith SD et al. Consistent breakage between consensus recombinase heptamers of chromosome 9 DNA in a recurrent chromosomal translocation of human T-cell leukemia. J Exp Med 1989; 169(2):369-377.
46. Cowell LG, Davila M, Yang K et al. Prospective estimation of recombination signal efficiency and identification of functional cryptic signals in the genome by statistical modeling. J Exp Med 2003; 197(2):207-220.
47. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. Cell 1985; 40(2):271-281.
48. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85(6):887-897.
49. Baumann M, Mamais A, McBlane F et al. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. EMBO J 2003; 22(19):5197-5207.
50. Muegge K, West M, Durum SK. Recombination sequence-binding protein in thymocytes undergoing T-cell receptor gene rearrangement. Proc Nat Acad Sci USA 1993; 90:4151-4155.
51. Kwon J, Imbalzano AN, Matthews A et al. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. Mol Cell 1998; 2(6):829-839.
52. Oettinger MA. How to keep V(D)J recombination under control. Immunol Rev 2004; 200:165-181.
53. Saha A, Wittmeyer J, Cairns BR. Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol 2006; 7(6):437-447.
54. Ruthenburg AJ, Li H, Patel DJ et al. Multivalent engagement of chromatin modifications by linked binding modules. Nat Rev Mol Cell Biol 2007; 8(12):983-994.
55. Li H, Ilin S, Wang W et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature 2006; 442(7098):91-95.
56. Pena PV, Davrazou F, Shi X et al. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature 2006; 442(7098):100-103.
57. Shi X, Hong T, Walter KL et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature 2006; 442(7098):96-99.
58. Wysocka J, Swigut T, Xiao H et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 2006; 442(7098):86-90.
59. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27(4):561-571.
60. Ramon-Maiques S, Kuo AJ, Carney D et al. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. Proc Natl Acad Sci USA 2007; 104(48):18993-18998.
61. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450(7172):1106-1110.
62. Krangel MS. T-cell development: better living through chromatin. Nat Immunol 2007; 8(7):687-694.
63. Hu H, Wang B, Borde M et al. Foxp1 is an essential transcriptional regulator of B-cell development. Nat Immunol 2006; 7(8):819-826.
64. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18(4):411-422.
65. Roldan E, Fuxa M, Chong W et al. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol 2005; 6(1):31-41.
66. Skok JA, Gisler R, Novatchkova M et al. Reversible contraction by looping of the Tcra and Tcrb loci in rearranging thymocytes. Nat Immunol 2007; 8(4):378-387.
67. Liang HE, Hsu LY, Cado D et al. The "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B- and T-cell development. Immunity 2002; 17(5):639-651.
68. Akamatsu Y, Monroe R, Dudley DD et al. Deletion of the RAG2 C terminus leads to impaired lymphoid development in mice. Proc Natl Acad Sci USA 2003; 100(3):1209-1214.
69. Steen SB, Gomelsky L, Speidel SL et al. Initiation of V(D)J recombination in vivo: role of recombination signal sequences in formation of single and paired double-strand breaks. EMBO Journal 1997; 16(10):2656-2664.
70. Bakhshi A, Wright JJ, Graninger W et al. Mechanism of the t(14; 18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. Proc Natl Acad Sci USA 1987; 84:2396-2400.

71. Posey JE, Brandt VL, Roth DB. Paradigm switching in the germinal center. Nat Immunol 2004; 5(5):476-477.
72. Weterings E, Chen DJ. The endless tale of nonhomologous end-joining. Cell Res 2008; 18(1):114-124.
73. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res 2008; 18(1):134-147.
74. Gottlieb TM, Jackson SP. The DNA-dependent protein kinase: Requirement for DNA ends and association with ku antigen. Cell 1993; 72:131-142.
75. Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 2001; 412(6847):607-614.
76. Roberts SA, Ramsden DA. Loading of the nonhomologous end joining factor, ku, on protein-occluded DNA ends. J Biol Chem 2007; 282(14):10605-10613.
77. Ma Y, Pannicke U, Schwarz K et al. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell 2002; 108(6):781-794.
78. Leber R, Wise TW, Mizuta R et al. The XRCC4 gene product is a target for and interacts with the DNA-dependent protein kinase. J Biol Chem 1998; 273(3):1794-1801.
79. Mahajan KN, Gangi-Peterson L, Sorscher DH et al. Association of terminal deoxynucleotidyl transferase with ku. Proc Natl Acad Sci USA 1999; 96(24):13926-13931.
80. Purugganan MM, Shah S, Kearney JF et al. Ku80 is required for addition of N nucleotides to V(D)J recombination junctions by terminal deoxynucleotidyl transferase. Nucleic Acids Res 2001; 29(7):1638-1646.
81. Critchlow SE, Bowater RP, Jackson SP. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. Current Biology 1997; 7:588-598.
82. Grawunder U, Wilm M, Wu X et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature 1997; 388:492-495.
83. Modesti M, Hesse JE, Gellert M. DNA binding of Xrcc4 protein is associated with V(D)J recombination but not with stimulation of DNA ligase IV activity. EMBO J 1999; 18(7):2008-2018.
84. Dai Y, Kysela B, Hanakahi LA et al. Nonhomologous end joining and V(D)J recombination require an additional factor. Proc Natl Acad Sci USA 2003; 100(5):2462-2467.
85. Buck D, Malivert L, de Chasseval R et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 2006; 124(2):287-299.
86. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 2006; 124(2):301-313.
87. Callebaut I, Malivert L, Fischer A et al. Cernunnos interacts with the XRCC4 × DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. J Biol Chem 2006; 281(20):13857-13860.
88. Tsai CJ, Kim SA, Chu G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. Proc Natl Acad Sci USA 2007; 104(19):7851-7856.
89. Lieber MR, Lu H, Gu J et al. Flexibility in the order of action and in the enzymology of the nuclease, polymerases and ligase of vertebrate nonhomologous DNA end joining: relevance to cancer, aging and the immune system. Cell Res 2008; 18(1):125-133.
90. Jhappan C, Morse HC, Fleischmann RD et al. DNA-PKcs: a T-cell tumour suppressor encoded at the mouse scid locus. Nat Genet 1997; 17:483-486.
91. Custer RP, Bosma GC, Bosma MJ. Severe combined immunodeficiency in the mouse: pathology, reconstitution, neoplasms. Am J Pathol 1985; 120:464-477.
92. Gu Y, Seidl KJ, Rathbun GA et al. Growth retardation and leaky SCID phenotype of Ku70-deficient mice. Immunity 1997; 7:653-665.
93. Li GC, Ouyang H, Li X et al. Ku70: a candidate tumor suppressor gene for murine T-cell lymphoma. Mol Cell 1998; 2(1):1-8.
94. Difilippantonio MJ, JZ, JT C et al. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. Nature 2000; 404:510-514.
95. Gao Y, Ferguson DO, WX et al. Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. Nature 2000; 404:897-900.
96. Zhu C, Mills KD, Ferguson DO et al. Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. Cell 2002; 109(7):811-821.
97. Moshous D, Pannetier C, de Chasseval R et al. Partial T and B lymphocyte immunodeficiency and predisposition to lymphoma in patients with hypomorphic mutations in Artemis. J Clin Invest 2003; 111:381-387.
98. Ferguson DO, Alt FW. DNA double strand break repair and chromosomal translocation: lessons from animal models. Oncogene 2001; 20(40):5572-5579.

99. Ferguson DO, Sekiguchi JM, Chang S et al. The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. Proc Natl Acad Sci USA 2000; 97(12):6630-6633.
100. Sharpless NE, Ferguson DO, O'Hagan RC et al. Impaired nonhomologous end-joining provokes soft tissue sarcomas harboring chromosomal translocations, amplifications and deletions. Mol Cell 2001; 8(6):1187-1196.
101. Gladdy RA, Taylor MD, Williams CJ et al. The RAG-1/2 endonuclease causes genomic instability and controls CNS complications of lymphoblastic leukemia in p53/Prkdc-deficient mice. Cancer Cell 2003; 3(1):37-50.
102. Riballo E, Critchlow SE, Teo S-H et al. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. Curr Biol 1999; 9:699-702.
103. Toita N, Hatano N, Ono S et al. Epstein-Barr virus-associated B-cell lymphoma in a patient with DNA ligase IV (LIG4) syndrome. Am J Med Genet A. 2007; 143(7):742-745.
104. Wilson JH, Berget PB, Pipas JM. Somatic cells efficiently join unrelated DNA segments end-to-end. Mol Cell Biol 1982; 2(10):1258-1269.
105. Roth DB, Porter TN, Wilson JH. Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol 1985; 5:2599-2607.
106. Roth DB, Wilson JH. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. Mol Cell Biol 1986; 6:4295-4304.
107. Alt FW, Baltimore D. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. Proc Natl Acad Sci USA 1982; 79:4118-4122.
108. Pergola F, Zdzienicka MZ, Lieber MR. V(D)J recombination in mammalian mutants defective in DNA double-strand break repair. Mol Cell Biol 1993; 13:3464-3471.
109. Taccioli GE, Rathbun G, Oltz E et al. Impairment of V(D)J recombination in double-strand break repair mutants. Science 1993; 260:207-210.
110. Getts RC, Stamato TD. Absence of a Ku-like DNA end binding activity in the xrs double- strand DNA repair-deficient mutant. J Biol Chem 1994; 269:15981-15984.
111. Taccioli GE, Cheng H-L, Varghese AJ et al. A DNA repair defect in chinese hamster ovary cells affects V(D)J recombination similarly to the murine scid mutation. J Biol Chem 1994; 269:7439-7442.
112. Taccioli GE, Gottlieb TM, Blunt T et al. Ku80: Product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. Science 1994; 265:1442-1445.
113. Kabotyanski EB, Gomelsky L, Han J-O et al. Double-strand break repair in Ku86- and XRCC4-deficient cells. Nucleic Acids Res 1998; 26(23):5333-5342.
114. Baumann P, West SC. DNA End-joining catalyzed by human cell-free extracts. Proc Natl Acad Sci USA 1998; 95:14066-14070.
115. Verkaik NS, Esveldt-van Lange RE, van Heemst D et al. Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. Eur J Immunol 2002; 32(3):701-709.
116. Lewis S, Gellert M. The mechanism of antigen receptor gene assembly. Cell 1989; 59:585-588.
117. Zhu C, Bogue MA, Lim D-S et al. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. Cell 1996; 86:379-389.
118. Huye LE, Purugganan MM, Jiang MM et al. Mutational analysis of all conserved basic amino acids in RAG-1 reveals catalytic, step arrest and joining-deficient mutants in the V(D)J recombinase. Mol Cell Biol 2002; 22(10):3460-3473.
119. Yarnall Schultz H, Landree MA, Qiu JX et al. Joining-deficient RAG1 mutants block V(D)J recombination in vivo and hairpin opening in vitro. Mol Cell 2001; 7(1):65-75.
120. Qiu JX, Kale SB, Yarnell Schultz H et al. Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol Cell 2001; 7(1):77-87.
121. Tsai CL, Drejer AH, Schatz DG. Evidence of a critical architectural function for the RAG proteins in end processing, protection and joining in V(D)J recombination. Genes Dev 2002; 16(15):1934-1949.
122. Richardson C, Jasin M. Coupled homologous and nonhomologous repair of a double-strand break preserves genomic integrity in mammalian cells. Mol Cell Biol 2000; 20(23):9068-9075.
123. Lee GS, Neiditch MB, Salus SS et al. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. Cell 2004; 117(2):171-184.
124. Corneo B, Wendland RL, Deriano L et al. Rag mutations reveal robust alternative end joining. Nature 2007; 449(7161):483-486.
125. Yan CT, Boboila C, Souza EK et al. IgH class switching and translocations use a robust nonclassical end-joining pathway. Nature 2007; 449(7161):478-482.
126. Soulas-Sprauel P, Le Guyader G, Rivera-Munoz P et al. Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. J Exp Med 2007; 204(7):1717-1727.

127. DiBiase SJ, Zeng ZC, Chen R et al. DNA-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. Cancer Res 2000; 60(5):1245-1253.
128. Meaburn KJ, Misteli T, Soutoglou E. Spatial genome organization in the formation of chromosomal translocations. Semin Cancer Biol 2007; 17(1):80-90.
129. Soutoglou E, Dorn JF, Sengupta K et al. Positional stability of single double-strand breaks in mammalian cells. Nat Cell Biol 2007; 9(6):675-682.
130. Lisby M, Mortensen UH, Rothstein R. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. Nat Cell Biol 2003; 5(6):572-577.
131. Lukasova E, Kozubek S, Kozubek M et al. Localisation and distance between ABL and BCR genes in interphase nuclei of bone marrow cells of control donors and patients with chronic myeloid leukaemia. Hum Genet 1997; 100(5-6):525-535.
132. Kozubek S, Lukasova E, Ryznar L et al. Distribution of ABL and BCR genes in cell nuclei of normal and irradiated lymphocytes. Blood 1997; 89(12):4537-4545.
133. Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol 2006; 4(5):e138.
134. Chen HT, Bhandoola A, Difilippantonio MJ et al. Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX. Science 2000; 290(5498):1962-1965.
135. Perkins EJ, Nair A, Cowley DO et al. Sensing of intermediates in V(D)J recombination by ATM. Genes Dev 2002; 16(2):159-164.
136. Bender CF, Sikes ML, Sullivan R et al. Cancer predisposition and hematopoietic failure in Rad50(S/S) mice. Genes Dev 2002; 16(17):2237-2251.
137. Celeste A, Petersen S, Romanienko PJ et al. Genomic instability in mice lacking histone H2AX. Science 2002; 296(5569):922-927.
138. Liyanage M, Weaver Z, Barlow C et al. Abnormal rearrangement within the alpha/delta T-cell receptor locus in lymphomas from Atm-deficient mice. Blood 2000; 96(5):1940-1946.
139. Kang J, Bronson RT, Xu Y. Targeted disruption of NBS1 reveals its roles in mouse development and DNA repair. EMBO J 2002; 21(6):1447-1455.
140. Difilippantonio S, Celeste A, Fernandez-Capetillo O et al. Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. Nat Cell Biol 2005; (7):675-685.
141. Ward IM, Difilippantonio S, Minn K et al. 53BP1 cooperates with p53 and functions as a haploinsufficient tumor suppressor in mice. Mol Cell Biol 2005; 25(22):10079-10086.
142. Liao MJ, Van Dyke T. Critical role for atm in suppressing V(D)J recombination-driven thymic lymphoma. Genes Dev 1999; 13(10):1246-1250.
143. Petiniot LK, Weaver Z, Vacchio M et al. RAG-mediated V(D)J recombination is not essential for tumorigenesis in atm-deficient mice. Mol Cell Biol 2002; 22(9):3174-3177.
144. Bredemeyer AL, Sharma GG, Huang CY et al. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature 2006; 442(7101):466-470.
145. Callen E, Jankovic M, Difilippantonio S et al. ATM prevents the persistence and propagation of chromosome breaks in lymphocytes. Cell 2007; 130(1):63-75.
146. Rosenwald A. DNA microarrays in lymphoid malignancies. Oncology (Williston Park). ec 2003; 17(12):1743-1748; discussion 1750, 1755, 1758-1749 passim.
147. Matsuoka S, Ballif BA, Smogorzewska A et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 2007; 316(5828):1160-1166.

Chapter 4

V(D)J Recombination Deficiencies

Jean-Pierre de Villartay*

Abstract

V(D)J recombination not only comprises the molecular mechanism that insures diversity of the immune system but also constitutes a critical checkpoint in the developmental program of B- and T-lymphocytes. The analysis of human patients with Severe Combined Immune Deficiency (SCID) has contributed to the understanding of the biochemistry of the V(D)J recombination reaction. The molecular study V(D)J recombination settings in humans, mice and in cellular mutants has allowed to unravel the process of Non Homologous End Joining (NHEJ), one of the key pathway that insure proper repair of DNA double strand breaks (dsb), whether they occur during V(D)J recombination or secondary to other DNA injuries. Two NHEJ factors, *Artemis* and *Cernunnos*, were indeed discovered through the study of human V(D)J recombination defective human SCID patients.

Introduction

Foreign antigens are recognized by the immune system of vertebrate through specialized receptors expressed on the cell surface of T- and B- lymphocytes; the T-cell receptors (TCR) and the B-cell receptors (BCR) or immunoglobulins respectively.

V(D)J Recombination

Immunoglobulins and TCR chains are composed of two domains: one constant region, which insures effector function and one highly polymorphic antigen recognition domain, or Variable domain. The Variable domain can be further subdivided into three separate segments known as Variable (V), Diversity (D) and joining (J) elements, whose respective encoding genes are dispersed on the chromosome (Fig. 1A). The fusion of these various elements, at the DNA level, by a site specific rearrangement process results in the formation of a functional V(D)J gene unit that will encode the Variable domain.[1] The combinatorial association of V, D and J elements thus enforces the required diversity of antigenic receptors. The V(D)J recombination reaction (Fig. 1B) is initiated by the lymphoid specific factors Rag1 and Rag2,[2,3] which specifically recognize recombination signal sequences (RSS) that flank all V, D and J gene units and introduce a DNA-dsb at the border of the RSS.[4] The resulting DNA-dsb is resolved by the ubiquitous DNA repair machinery known as nonhomologous end joining (NHEJ). As discussed below, the V(D)J recombination process not only enforces the diversity of the immune system, it also can be considered as a critical checkpoint in the development of B- and T-lymphocytes as a faulty V(D)J reaction leads to an arrest in the differentiation of these two lineages (Fig. 2) causing a Severe Combined Immune Deficiency (SCID) phenotype.

*Jean-Pierre de Villartay—INSERM U768, Hôpital Necker Enfants Malades, 149 rue de Sèvres, 75015 Paris, France. Email: devillar@necker.fr

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

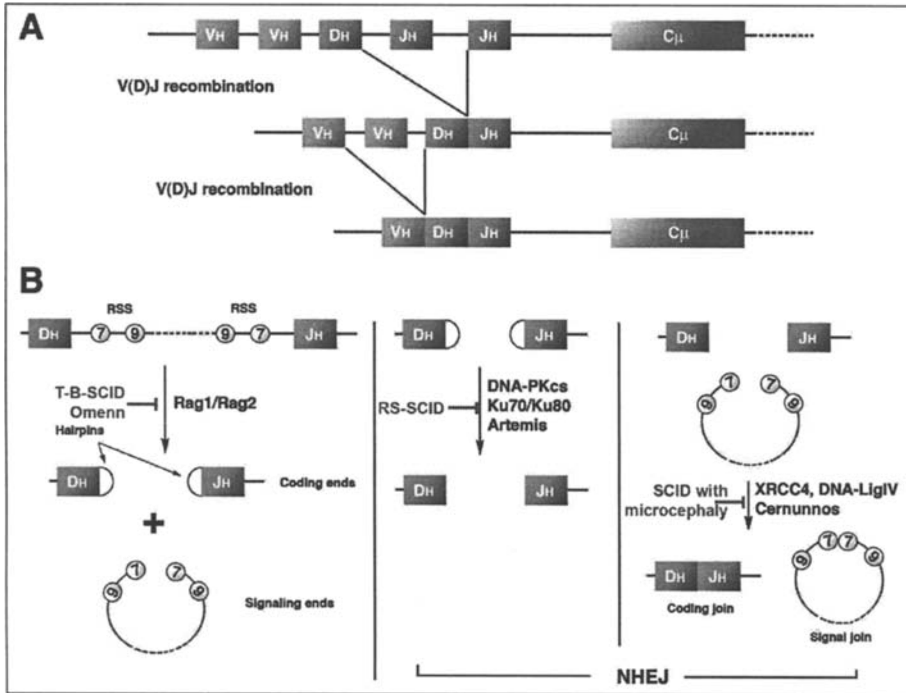


Figure 1. V(D)J recombination. A) Organization of IgH locus and rearrangement process. B) The V(D)J recombination reaction and faulty steps in SCIDs.

Human Primary Immune Deficiencies

Severe combined immune deficiencies (SCIDs) comprise about eleven inherited rare disorders which have in common a block of T-cell differentiation/function associated with a direct or indirect impairment of B-cell immunity.[5] As a consequence of their molecular defects, the clinical presentation of SCID patients is rather uniform and mainly characterized by the early onset of infections affecting the respiratory tract and the gut. Common opportunistic organisms (*Pneumocystis carinii* and *Aspergillus*) as well as viruses (*Cytomegalovirus* for example) cause recurrent infections and failure to thrive. About 30% of human SCID cases arise from a defect in V(D)J recombination (Fig. 2, Table 1), leading to an early arrest of both B- and T-lymphocyte maturation. This T-B-SCID condition can be either the result of deleterious mutations in the *Rag1* and *Rag2* genes[6] affecting the initiation of the V(D)J recombination, or impinge on the DNA repair phase of the V(D)J recombination reaction. In the latter case the immune deficiency is accompanied by an increased cellular sensitivity to ionizing radiation (RS-SCIDs), a condition resembling the murine scid situation. In addition to these rather straightforward clinical presentations, several other immune deficiencies caused by variable defects in V(D)J recombination have been described more recently, which are associated with additional developmental anomalies such as a facial dismorphy and microcephaly. The molecular analyses of these human and mouse pathologies were highly instrumental in defining some of the actors of the nonhomologous end-joining pathway.

RAG1 and *RAG2* Deficiencies

T-B-SCIDs

The first evidence for a critical role of V(D)J recombination for the proper development of both B- and T-lymphocytes came from the analysis of *Rag1* and *Rag2* KO mice.[7,8] Both mouse strains

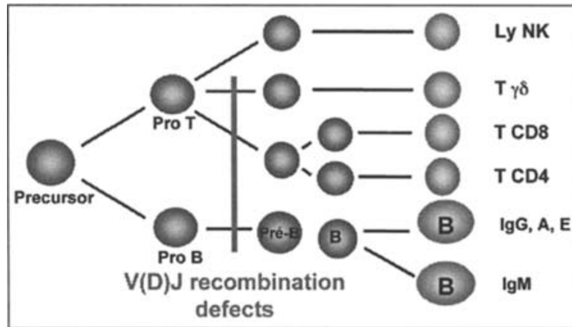


Figure 2. Specific block of B- and T-cell maturation in V(D)J recombination deficiencies.

are characterized by a complete absence of peripheral mature B- and T-lymphocytes owing to a defect in the initiation of the V(D)J recombination. *Rag1* and *Rag2* deficiency thymocytes accumulate as quiescent cells at the CD4-CD8- double negative (DN) stage, just prior the onset of TCR-β rearrangement. Similarly, B-cells development is arrested at the preB-cell stage in the bone marrow. Except for their immunological phenotype, *Rag1* and *Rag2* KO mice do not develop any other functional anomaly. A V(D)J recombination defect was subsequently identified in a group of human patients presenting with the same clinical/biological condition,[9,10] which turned out to be caused by mutations in either one of the Rag genes.[11] The *Rag1* and *Rag2* deficiencies, MIM (mendelian Inheritance in Man) #179615 and #179616 respectively, are autosomal recessive diseases. Both genes are located on human chromosome 11p13 and carriers of heterozygous mutations are healthy without any immunological disturbance. Apart from the finding of a complete alymphocytosis in the blood, no simple functional assays are available to reveal a V(D)J recombination defect caused by *Rag1* or *Rag2* mutations.

Rag1/2 Structure and Function

The biochemistry of the initial steps of the V(D)J recombination and the precise function of the *Rag1*/2 proteins are detailed in other chapters of this book. The identification of a whole series of mutations in either genes from the molecular analysis of human T-B-SCID patients over the years was highly instrumental in drawing structure/function relationships that help defining the various functional domains of these two proteins.[12]

Omenn Syndrome

Omenn syndrome (OS) was first described in 1965 as a rare autosomal recessive disease (MIM #2603554) characterized by an immunodeficiency accompanied by a severe erythroderma caused by skin infiltrating activated lymphocytes (Fig. 3), eosinophilia, hepatosplenomegaly, lymphadenopathy, high level of IgE but very low levels of the other Ig isotypes.[13] The existence of both T-B-SCID and Omenn syndrome in the same family[14] suggested that OS could result from a V(D)J recombination defect caused by *Rag1*/2 mutations. Mutations in the *Rag1* and *Rag2* genes were indeed identified in several cases of OS.[12,15,16] These mutations are by essence hypomorphic as they allow V(D)J recombination and hence the development of B- and T-lymphocytes, to proceed to some extent. Consistent with the observation that OS appears in the context of a faulty, although not complete failure, of T- and sometimes B-cell development, particular mutations in *Artemis* (MIM #605988),[17] the α chain of the IL7 receptor (MIM #600802)[18] and in the Mitochondrial RNA-processing endoribonuclease (RMRP; MIM #157660) gene[19] were reported to cause OS. Lastly, OS like phenotype was also noted in DiGeorge syndrome.[20] Conversely, leaky V(D)J recombination is not always associated with the development of OS[12] and recent reports identified hypomorphic *Rag1* and *Rag2* mutations in patients characterized by an elevated count of TCR-γ/δ expressing T-lymphocytes in the peripheral blood secondary to CMV infection in the

Table 1. Gene defects in T-B- SCID and RS SCIDs

Gene	Mutation	Radiosensitivity	Immune Defect	Growth Delay	Microcephaly	Cancer Predisposition	Ref
Rag1, Rag2	Null	No	T-B-SCID	No	No	No	11
	Hypomorphic	No	Omenn	No	No	No	15
Artemis	Null	Yes	RS-SCID	No	No	?	45
	Hypomorphic	Yes	RS-SCID	No	No	Yes	107
	Hypomorphic	Yes	Omenn	No	No	No	17
DNA-LigIV	Hypomorphic	Yes	μ-SCID	Yes	Variable	Yes	77
Cernunnos	Hypomorphic?	Yes	μ-SCID	Yes	Yes	?	82

context of almost complete T-B-SCID.[21,22] Altogether, these reports suggest that additional genetic or environmental factors may be required for the oligoclonal expansion of few T-cell clones that emerge as a result of drastically reduced output of T-cells.

Physiopathology

One characteristic feature of OS is the infiltration of the skin and gastrointestinal mucosa with activated T-lymphocytes causing a skin rash resembling that occurring during the graft versus host reaction (GVH). Activated T-lymphocytes with a highly restricted TCR heterogeneity are also present in the blood of OS patients.[14] These activated T-cells, which are skewed toward the TH2 phenotype, are probably responsible for the associated eosinophilia and hyper IgE secretion.[23] Such an autoimmune like disease strongly suggests that T-cell tolerance does not occur properly in this condition. A recent study has tackle the idea that AIRE deficiency in OS could be at the base of the autoimmune manifestations.[24] AIRE (Autoimmune regulator element) is a transcriptional activator expressed by medullary epithelial cells in the thymus. AIRE regulates the ectopic expression of a set of tissue-specific proteins, which are normally expressed in the periphery, thus driving central tolerance towards these proteins.[25] Mutations of AIRE cause autoimmune polyendocrynopathy-candidiasis-ectodermal dystrophy (APECED). Since AIRE expression in the thymus requires normal T-cell development, it was proposed that thymic AIRE expression in OS could somehow be impaired resulting in faulty negative selection and survival of autoimmune T-lymphocytes. Indeed two thymi from OS patients showed a strongly reduced AIRE expression both at the protein and mRNA level.[24] Although this finding provides an important breakthrough in the understanding of OS physiopathology, several questions remain unanswered. The recent design of animal models mimicking some of the key OS features will clarify some of these questions.

Murine Models of Omenn Syndrome

Marrella et al. Introduced the *Rag2* R229Q mutation, known to be associated with OS and atypical SCID in humans, by means of knock-in on a full *Rag2* KO background.[26] About half of the resulting mice developed alopecia, erythroderma, wasting syndrome and colitis by the age of three months. This phenotype, resembling human OS, is also accompanied with T-cell infiltration of the skin and the gut. These mice also experienced a severe depletion of their B-cell count in spleen and a partial block of thymocyte development at the DN stage, prior to TCR-β rearrangements. Like

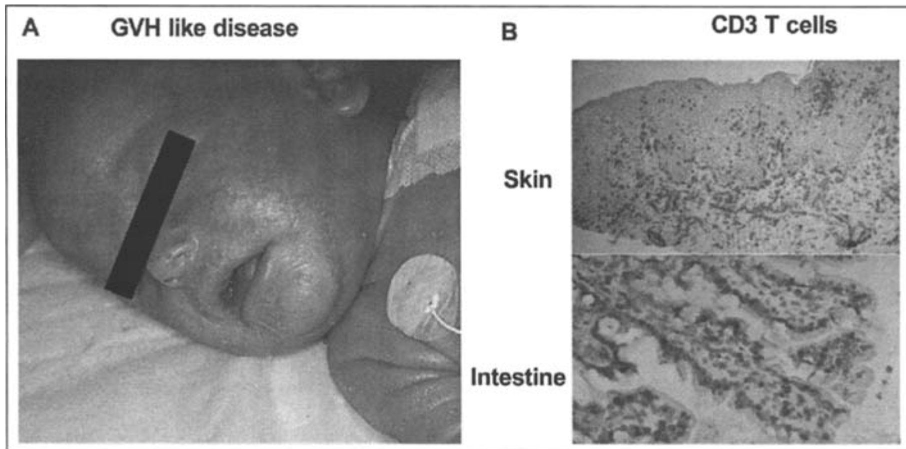


Figure 3. GVH like disease in Omenn syndrome. A) Skin erythroderma B) T-cell infiltration of the skin and the intestine.

in the human OS condition, the thymic expression of AIRE is severely reduced in these mice. Interestingly, the two particular sub population of T-lymphocytes regulatory T-cells (Treg) and invariant natural killer cells (iNKT) are also strongly reduced in *Rag2*^{R229Q} mice. The absence of iNKT cells in human OS was recently demonstrated and may well participate in the physiopathology of OS condition.[27] During a systematic survey of their mouse colony for the appearance of anomaly in the development of T-cells, Khiong et al identified a spontaneous mutant mouse with an increased number of so called "memory" T-cells.[28] It turned out that these mice carried a *Rag1* R972G mutation and presented with many characteristic features of human OS in whom the equivalent *Rag1* R975 amino acid was found mutated. Based on the introduction of this mutation into a CD4 KO background, the authors propose that the abnormal homeostasis of CD4+ T-cells could participate in the onset of OS manifestations. These two OS mouse models will certainly be very helpful in the future for the better understanding of the OS condition, in particular with regard to the possible impact of environmental factors on the development of the autoimmune manifestations. A third interesting model of OS came out from studies in the WE Paul laboratory. Milner and col. showed that reconstitution of lymphopenic (*Rag2* KO) mice with suboptimal numbers of T-lymphocytes results in a multiorgan inflammatory disease resembling OS.[29,30] These authors show that reconstitution of *Rag2* KO mice with a small number of T-cells , in contrast to a large number of T-cells, results in the onset of OS like phenotype. Indeed it is not the absolute number of T-cells per se that causes this phenotype but rather the reduced TCR diversity of the transplanted T-cells. They could further link this phenotype to the reduced heterogeneity of the TCR repertoire expressed by the limited numbers of Treg in their inoculums. This study very nicely complements the data obtained with the two OS mice and the observation gathered from human OS condition.

T-B-SCID with Radiosensitivity

The Scid Mouse and the CHO-XRCC Mutants

The description of the scid mouse,[31] a natural mutant mouse characterized by a lack of circulating B- and T-lymphocytes, as resulting from a general DNA repair defect accompanied by an increased sensitivity to ionizing radiation or other agents causing DNA dsb, provided the initial link between V(D)J recombination and general DNA-dsb repair.[32-34] The faulty V(D)J recombination in scid mice can be demonstrated in pre-T and pre-B-cells using extrachromosomal V(D)J recombination substrates[35] as well as on endogenous TCR loci in ex vivo isolated thymocytes.[36]

The design of V(D)J recombination substrates was at the base of the strategy for the identification of the *Rag1* and *Rag2*[2,3] genes and are still in use in many laboratories to assess various aspects of V(D)J recombination. Another major breakthrough came from the very clever idea of performing V(D)J assays in mutagenized Chinese hamster ovary cells (CHO) that had been initially selected for their DNA repair defect revealed by an increased X ray sensitivity (XRCC mutants). Several of these mutants happened to be V(D)J recombination defective.[37,38] To summarize years of intensive work in many laboratories, these experiments entitled the definition of two important protein complexes at play during NHEJ; The DNA-PK complex formed by Ku70 (XRCC6), Ku80 (XRCC5) and the DNA-PKcs catalytic subunit (XRCC7, Mu-scid) on one hand and the XRCC4/DNA-LigaseIV on the other hand. The precise function and activities of these NHEJ factors have been thoroughly reviewed in recent years.[39] Briefly, the DNA-PK complex identifies the Rag1/2 generated DNA-dsb (Fig. 1B) and recruits the processing enzyme Artemis (see below) while the XRCC4/DNA-LigaseIV complex, together with Cernunnos (see below), terminates the reaction by rejoining the broken DNA ends.

Artemis

Some B-T-SCID patients do not harbor mutation in either *Rag1* or *Rag2* genes, yet they present the same clinical/biological features as *Rag1/2* defective T-B- SCIDs, i.e., a complete absence of circulating mature B- and T-lymphocytes. The alymphocytosis in these patients is accompanied by an increased sensitivity to ionizing radiations of bone marrow cells (CFU-GM) and skin fibroblasts.[40] This characteristic, also shared by the scid mice, led to the hypothesis of a general DNA repair defect in RS-SCID patients. The RS-SCID phenotype is also found with high incidence among Athabascan-speaking Native American Indians. Consistent with their general DNA repair deficiency, they present a V(D)J recombination defect which can be demonstrated in vitro, in fibroblasts, using V(D)J recombination substrates and ectopic expression of both *Rag1* and *Rag2* genes.[41,42] Despite the strong similarity of RS-SCID patients and scid mice, DNA-PK activity is normal in these patients and the implication of the *DNA-PKcs* gene has been ruled out by genetic means in several families.[43] The disease-related locus in RS-SCID was assigned to the short arm of the chromosome 10.[41,44] Given the location of the RS-SCID locus on human chromosome 10, genomic DNA sequences covering this region were analyzed in silico for the presence of putative genes, leading to the identification of a new DNA repair factor called Artemis.[45] Functional complementation studies and mutation analyses certified that *Artemis* was indeed the gene defective in RS-SCID. Consistent with its function during V(D)J recombination and DNA repair, *Artemis* is ubiquitously expressed and is localized in the nucleus. *Artemis* mutations, which account for the RS-SCID condition, are primarily localized in the N-terminus half of the protein, thought to harbor the catalytic domain. These mutations involve nonsense and misense substitutions as well as splicing defects leading to severely truncated proteins. The inactivation of the *Artemis* gene in mice recapitulates the clinical and biological features of RS-SCID patients.[46,47] Hypomorphic *Artemis* mutations have been identified in patients presenting a leaky SCID phenotype[17,48] as well as in one patient characterized by a progressive combined immune deficiency resulting from an elevated lymphocyte apoptosis but a delayed cell death of IR treated fibroblasts in vitro.[49]

Artemis Structure and Function

In depth in silico *Artemis* sequence analysis highlighted significant similarities of the first 150 amino acids to well-established members of the metallo-β-lactamase superfamily.[45] The metallo-β-lactamase fold is adopted by various metallo-enzymes with a widespread distribution and substrate specificity.[50] It consists of a four-layered β sandwich with two mixed ß sheets flanked by α helices. Biochemical studies demonstrated that Artemis does indeed exert an intrinsic 5′ to 3′ exonuclease activity in vitro.[51] A similar exonuclease activity has also been recognized in Apollo/SNM1B, a protein related to Artemis that functions in the protection of telomeres.[52,53] When Artemis is associated with and phosphorylated by DNA-PKcs it switches its catalytic activity to a DNA endonuclease capable of opening Rag1/2 generated hairpin structures during V(D)J recombination.[51] Consistent with this hairpin opening activity, Artemis

and DNA-PKcs are required for efficient adeno-associated virus (AAV) infection, the process of which goes through resolution of hairpin loops at the AAV inverted terminal repeat (ITR) extremity of the viral DNA genome.[54] Sequence analysis, secondary structure prediction and mutagenesis studies clearly indicated the conservation of motifs (HxHxDH) typical of the metallo-β-lactamase fold, participating in the metal binding pocket and representing the catalytic site of the metallo-β-lactamases.[55-58] Following the metallo-β-lactamase domain, Artemis shares several conserved features with other metallo-β-lactamases acting specifically on nucleic acids and involved in DNA repair (Artemis, SNM1, PSO2) and RNA processing (CPSF). This new domain was called β-CASP.[55] The β-CASP domain, which is always appended to a metallo-β-lactamase domain, is strictly required for Artemis function.[56] The three-dimensional structure of several RNA-specific β-CASP members has recently revealed the general organization of these proteins into two domains: a metallo-β-lactamase domain and a β-CASP domain, with the active site being located at the interface between the two domains.[59,60] Several Serine residues, mostly located in the C-terminus half of the protein, have been identified in vitro and in vivo as targets of phosphatidylinositol 3- kinase like kinases (PIKK), including DNA-PKcs.[61-68] Unexpectedly however, Artemis function in V(D)J recombination does not rely upon the phosphorylation of these sites. Moreover, a truncated version of Artemis lacking the C-terminus half is still proficient in V(D)J recombination.[58,61] One current hypothesis is that, in the absence of DNA-PKcsArtemis would adopt a particular conformation by which its C-terminus domain masks the β-Lact/β-CASP catalytic site. Artemis would then gain its full enzymatic activity through conformational changes upon DNA-PKcs interaction.[56,65] Another proposed function for DNA-PKcs would be to facilitate the access of Artemis to DNA damage. DNA-PKcs is indeed required for the proper loading of Artemis on damaged chromatin.[69] However, although DNA-PK kinase activity prevents Artemis dissociation from the DNA-PK/DNA complex, it is the autophosphorylation of DNA-PKcs and not that of Artemis which is critical for the ultimate activation of Artemis endonuclease activity,[68] which suggests that conformational changes triggered by DNA-PKcs autophosphorylation expose DNA ends for further processing by Artemis.

Artemis and the DNA Damage Response

RS-SCID patients and Artemis KO mice present, in addition to their V(D)J recombination defect, a general increased cellular sensitivity to DNA damaging agents, arguing for an Artemis function during the repair of these damages. Indeed, the repair of about 10% of DNA lesions inflicted by ionizing radiations rely on Artemis as shown by the retention of γH2AX foci, a marker of DNA breaks,[70] on a fraction of cells at late time points following IR.[62,71] Artemis was found to process 3′-phosphoglycolate terminally blocked DSB in vitro, DNA modifications known to be induced by IR or bleomycin in vivo.[72] Artemis thus appears to be one constituent of the DNA damage response (DDR). The DDR is orchestrated by a series of biochemical events among which protein phosphorylation by the PIKK kinases, ATM and ATR, play a central role.[73] Like many DNA repair factors, Artemis is hyperphosphorylated in an ATM dependent manner after IR.[61-66,68,74] The exact role of ATM dependent phosphorylation of Artemis during DNA repair is not fully understood as mutations of the posphorylation sites do not impact on the capacity of these Artemis mutants to complement the radiosensitivity of Artemis deficient fibroblasts.[61] In addition to the DNA repair per se, cell cycle checkpoints constitute another key feature of the DDR. Following DNA damage, the cells arrest their cycling at the G1/S and the G2/M boundaries to allow DNA repair to proceed. In the case of IR induced DNA damage, these cell cycle checkpoints depend on ATM. Whether Artemis participates in cell cycle checkpoints remained a matter of debate. Although it is clear that Artemis deficient cells arrest normally in G1 following IR, the maintenance and/or recovery from the G2/M checkpoint following IR was found altered.[63,66,75,76] Whether this reflects a direct function of Artemis on cell cycle through the regulation of Cdk1-cyclin B[63] or the impaired repair of a subset of damage after IR[75,76] remains an open issue.

DNA-LigaseIV

DNA-LigaseIV mutations were first identified in patients presenting developmental anomalies and immunodeficiency.[77] In contrast to RS-SCIDs these patients are not completely devoid of B- and T-lymphocytes, although their numbers can be drastically reduced. Several other reports of *DNA-LigaseIV* deficiency further demonstrated the high heterogeneity of this syndrome for its impact on immunodeficiency (from no deficiency to SCID) as well as on its developmental consequences (with or without microcephaly) and cancer incidence.[78-81] In the more severe forms, the V(D)J recombination is strongly affected both quantitatively and qualitatively as a consequence of the DNA rejoining deficiency. Whatever the nature of human *DNA-LigaseIV* mutations, they all result in partial loss of function alleles.

Cernunnos

Another series of five patients characterized by severe combined immunodeficiency associated with growth delay and microcephaly was reported.[82] The clinical and cellular phenotypes of these patients (including increased radiosensitivity, defective V(D)J recombination, impaired in vitro NHEJ activity) was strikingly reminiscent to that observed in *DNA-LigaseIV* condition (see above). However, neither *DNA-LigaseIV* nor the other known NHEJ factors were found mutated, suggesting that these patients suffered from a novel NHEJ defect.[83] A new NHEJ factor, named Cernunnos, was indeed identified through cDNA functional complementation of patients' fibroblasts. The same NHEJ factor, named XLF (for XRCC4-like factor), was independently identified through a yeast two hybrid screen using XRCC4 as a bait.[84] Recently developed murine Cernunnos-deficient ES cells present a phenotype similar to that of human deficient cells (increased radiosensitivity, genomic instability, DNA repair defect), except for V(D)J recombination.[85] Although the efficiency of V(D)J recombination is highly compromised, the fidelity of signal joins is not altered in Cernunnos-deficient ES cells, contrasting with the human situation from which more than half of the signal joins are imprecise, with various lengths of nucleotide deletions.[82,86] The nature of the mutation engineered in ES cells (the deletion of *Cernunnos* exons 4 and 5 could result in the low level expression of a truncated Cernunnos protein created by an in-frame splicing from exon 3 to exon 6) may partly account for these differences.[85]

Deleterious mutations of the Cernunnos gene were found in all patients and the ectopic expression of a wild type Cernunnos complemented the DNA repair defect observed in patients' cells.[82,84] Whether these mutations lead to a complete loss of function or represent hypomorphic alleles is not yet known with certainty. Given the structural and functional relationships between Cernunnos/XLF and XRCC4 (see below), one would expect a complete loss of function allele not to be compatible with life as is the case for XRCC4 KO mice. The development of a Cernunnos complete loss of function mouse model will certainly help to address these issues.

Cernunnos Structure

The human *Cernunnos* gene, composed of eight exons, is located on the long arm of chromosome 2 (2q35) and is expressed as a 2063 nucleotides long cDNA.[82,84] The Cernunnos/XLF protein is 299 amino acid long with an apparent weight of about 33kDa. Cernunnos is ubiquitously expressed and localized predominantly in the nucleus. Sequence analysis revealed that Cernunnos shares structural features with XRCC4 revealing the existence of a new protein family.[55,84,87] Based on the XRCC4 structure[88,89] one can predict a similar conformation for Cernunnos, i.e., a globular head domain followed by a coil-coiled tail.[84,87,90] Cernunnos/XLF, like XRCC4, can bind DNA in a sequence-independent manner.[87,91] Cernunnos/XLF and XRCC4 can homodimerize or participate in the same complex together with DNA-LigaseIV.[84,87,90,92] Their globular head domains could drive their direct association. Both Cernunnos/XLF and XRCC4 appear to directly interact with DNA-LigaseIV but the Cernunnos/XLF-DNA-LigaseIV interaction is very weak.[87,92] The exact nature of the complex(es) formed between XRCC4, DNA-LigaseIV and Cernunnos/XLF remains to be clearly established, but one can anticipate that differential

complex formation may have important regulatory function for the DNA-end ligation reaction during the NHEJ process.

Lastly, sequence analysis revealed that Cernunnos/XLF, although highly divergent, is the genuine ortholog of Nej1p/Lif2,[90] a NHEJ factor described in the yeast *S. cerevisiae*.[93-95] Cernunnos orthologs (referenced as Nej1p or XLF1) have further been found in many eukaryotes demonstrating that Nej1p and Cernunnos/XLF belong to the same protein family.[87,90,96] Nej1p in yeast interacts with the XRCC4 ortholog Lif1p, suggesting that Nej1p and Cernunnos/XLF have conserved an analogous function throughout evolution.

Cernunnos Function

Like XRCC4 and several other factors that participate in the DNA damage response (DDR), Cernunnos/XLF and its yeast ortholog Nej1p are phosphorylated upon DNA damage.[97,98] However, the recruitment of Cernunnos to the site of DNA breaks does not require this DNA-PK dependent phosphorylation event.[97] Although XRCC4 and Cernunnos share structural characteristics and are part of the same complex, the over expression of XRCC4 cannot functionally complement Cernunnos deficient cells,[90] suggesting that these two factors participate to the DNA-end ligation activity in a cooperative manner. Moreover, the defects of XRCC4 or Cernunnos have different impact on the DNA-LigaseIV protein. Whereas DNA-LigaseIV protein is destabilized in the absence of XRCC4,[99,100] this is not the case in Cernunnos deficient cells.[84,90] Although the XRCC4/DNA-LigaseIV complex exerts DNA-end ligation in vitro,[101] Cernunnos/XLF further potentiates this activity.[87,91] The presence of Cernunnos, which seems particularly important for the ligation of mismatched or non cohesive DNA ends but not of compatible DNA ends in vitro[102,103] would suggest that it may potentiate the ligation activity of the XRCC4/DNA-LigIV complex on specific DNA end structures. Although the information concerning the role of Cernunnos are still scarce, the attractive hypothesis that XRCC4 stabilizes DNA-LigaseIV whilst Cernunnos switches-on the ligase activity of the XRCC4/DNA-LigaseIV complex can however be drawn. Hence, several corollaries follow this hypothesis: (1) Cernunnos might be a crucial regulator of the NHEJ process (as is the case for its S. cerevisiae ortholog Nej1p, see below) and (2) The Cernunnos ability to interact with the DNA-LigIV/XRCC4 complex and/or to associate with the DNA breaks and/or to potentiate the ligase activity should be tightly regulated (either transcriptionally as is the case for Nej1p, or posttranscriptionally or both). These hypotheses will be certainly tested in the next future and the structural analysis of Cernunnos crystal alone or in association with XRCC4 and DNA-LigaseIV will also be of great interest to unravel the specific role of *Cernunnos*.

V(D)J Recombination in NHEJ Deficient Animal Models

In addition to the scid mouse, deficient animal models were developed for the various NHEJ factors. All these models have in common an impact on V(D)J recombination and consequently on lymphocyte developmental arrest, thus recapitulating the human RS-SCID condition.[83] In the case of *XRCC4* and *DNA-LigaseIV* KO mice, the immunological phenotype is accompanied by embryonic lethality owing to a massive apoptosis of postmitotic neurons,[104,105] the corollary of which in humans could be the microcephaly observed in *DNA-LigaseIV* and *Cernunnos* patients. Another very interesting aspect came out from the analyses of these models. When the NHEJ defect is crossed onto a *P53* KO background, this invariably leads to the early onset of very aggressive Pro-B-cell lymphomas bearing chromosomal translocations, thus demonstrating that NHEJ factors are genetic caretakers.[106]

Acknowledgements

I thank all present and past members of my team who worked on the molecular basis of human SCID cases over the years as well as the colleagues in our Institute. Our work is supported by grants from Institut National de la Santé et de la Recherche Médicale (INSERM), the Agence Nationale de la Recherche (ANR), the Ligue Nationale contre le Cancer (Equipe labellisée LA LIGUE 2005), The Commissariat à l'Energie Atomique (LRC-CEA N°40V) and the INCa/Cancéroplole IdF.

References

1. Tonegawa S. Somatic generation of antibody diversity. Nature 1983;302(5909):575-81.
2. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell 1989; 59(6):1035-48.
3. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248(4962):1517-23.
4. Dudley DD, Chaudhuri J, Bassing CH et al. Mechanism and Control of V(D)J Recombination versus Class Switch Recombination: Similarities and Differences. Adv Immunol 2005; 8643-112.
5. Fischer, A. Human primary immunodeficiency diseases: a perspective. Nat Immunol 2004; 5(1):23-30.
6. Villa A, Sobacchi C, Notarangelo LD et al. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. Blood 2001; 97(1):81-8.
7. Shinkai Y, Rathbun G, Lam KP et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell 1992; 68(5):855-67.
8. Mombaerts P, Iacomini J, Johnson RS et al. RAG-1-deficient mice have no mature B- and T-lymphocytes. Cell 1992; 68(5):869-77.
9. Schwarz K, Hansen-Hagge TE, Knobloch C et al. Severe combined immunodeficiency (SCID) in man: B-cell-negative (B-) SCID patients exhibit an irregular recombination pattern at the JH locus. J Exp Med 1991; 174(5):1039-48.
10. Abe T, Tsuge I, Kamachi Y et al. Evidence for defects in V(D)J rearrangements in patients with severe combined immunodeficiency. J Immunol 1994; 152(11):5504-13.
11. Schwarz K, Gauss GH, Ludwig L et al. RAG mutations in human B-cell-negative SCID. Science 1996; 274(5284):97-9.
12. Sobacchi C, Marrella V, Rucci F et al. RAG-dependent primary immunodeficiencies. Hum Mutat 2006; 27(12):1174-84.
13. Omenn GS. Familial Reticuloendotheliosis with Eosinophilia. N Engl J Med 1965; 273:427-32.
14. de Saint-Basile G, Le Deist F, de Villartay JP et al. Restricted heterogeneity of T-lymphocytes in combined immunodeficiency with hypereosinophilia (Omenn's syndrome). J Clin Invest 1991; 87(4):1352-9.
15. Villa A, Santagata S, Bozzi F et al. Partial V(D)J recombination activity leads to Omenn syndrome. Cell 1998; 93(5):885-96.
16. Corneo B, Moshous D, Gungor T et al. Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. Blood 2001; 97(9):2772-6.
17. Ege M, Ma Y, Manfras B et al. Omenn syndrome due to ARTEMIS mutations. Blood 2005; 105(11):4179-86.
18. Giliani S, Bonfim C, de Saint Basile G et al. Omenn syndrome in an infant with IL7RA gene mutation. J Pediatr 2006; 148(2):272-4.
19. Roifman CM, Gu Y, Cohen A. Mutations in the RNA component of RNase mitochondrial RNA processing might cause Omenn syndrome. J Allergy Clin Immunol 2006; 117(4):897-903.
20. Markert ML, Alexieff MJ, Li J et al. Complete DiGeorge syndrome: development of rash, lymphadenopathy and oligoclonal T-cells in 5 cases. J Allergy Clin Immunol 2004; 113(4):734-41.
21. Ehl S, Schwarz K, Enders A et al. A variant of SCID with specific immune responses and predominance of gamma delta T-cells. J Clin Invest 2005; 115(11):3140-8.
22. de Villartay JP, Lim A, Al-Mousa H et al. A novel immunodeficiency associated with hypomorphic RAG1 mutations and CMV infection. J Clin Invest 2005; 115(11):3291-9.
23. Schandene L, Ferster A, Mascart-Lemone F et al. T-helper type 2-like cells and therapeutic effects of interferon-gamma in combined immunodeficiency with hypereosinophilia (Omenn's syndrome). Eur J Immunol 1993; 23(1):56-60.
24. Cavadini P, Vermi W, Facchetti F et al. AIRE deficiency in thymus of 2 patients with Omenn syndrome. J Clin Invest 2005; 115(3):728-32.
25. Anderson MS, Venanzi ES, Klein L et al. Projection of an immunological self shadow within the thymus by the aire protein. Science 2002; 298(5597):1395-401.
26. Marrella V, Poliani PL, Casati A et al. A hypomorphic R229Q Rag2 mouse mutant recapitulates human Omenn syndrome. J Clin Invest 2007; 117(5):1260-9.
27. Matangkasombut P, Pichavant M, Saez DE et al. Lack of iNKT cells in patients with combined immune deficiency due to hypomorphic RAG mutations. Blood 2007 [epub ahead of print.]
28. Khiong K, Murakami M, Kitabayashi C et al. Homeostatically proliferating CD4 T-cells are involved in the pathogenesis of an Omenn syndrome murine model. J Clin Invest 2007; 117(5):1270-81.
29. Milner JD, Ward JM, Keane-Myers A et al. Lymphopenic mice reconstituted with limited repertoire T-cells develop severe, multiorgan, Th2-associated inflammatory disease. Proc Natl Acad Sci USA 2007; 104(2):576-81.

30. Milner J, Ward J, Keane-Myers A et al. Repertoire-dependent immunopathology. J Autoimmun 2007; 29 (4):257-61.
31. Bosma GC, Custer RP, Bosma MJ. A severe combined immunodeficiency mutation in the mouse. Nature 1983; 301(5900):527-30.
32. Fulop GM, Phillips RA. The scid mutation in mice causes a general defect in DNA repair. Nature 1990; 347479-482.
33. Biedermann KA, Sum JR, Giaccia AJ et al. Scid mutation in mice confers hypersensitivity to ionizing radiation and a deficiency in DNA double-strand break repair. Proc Natl Acad Sci USA 1991; 88:1394-1397.
34. Hendrickson EA, Qin XQ, Bump EA et al. A link between double-strand break related repair and V(D)J recombination : the scid mutation. Proc Natl Acad Sci USA 1991; 88:4061-4065.
35. Lieber MR, Hesse JE, Lewis S et al. The defect in murine severe combined immune deficiency: joining of signal sequences but not coding segments in V(D)J recombination. Cell 1988; 55(1):7-16.
36. Roth DB, Menetski JP, Nakajima PB et al. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell 1992; 70(6):983-91.
37. Taccioli GE, Rathbun G, Oltz E et al. Impairment of V(D)J recombination in double-strand break repair mutants. Science 1993; 260(5105):207-10.
38. Pergola F, Zdzienicka MZ, Lieber MR. V(D)J recombination in mammalian cell mutants defective in DNA double-strand break repair. Mol Cell Biol 1993; 13(6):3464-71.
39. Weterings E, van Gent DC. The mechanism of nonhomologous end-joining: a synopsis of synapsis. DNA Repair (Amst) 2004; 3(11):1425-35.
40. Cavazzana-Calvo M, Le Deist F, de Saint Basile G et al. Increased radiosensitivity of granulocyte macrophage colony-forming units and skin fibroblasts in human autosomal recessive Severe Combined Immunodeficiency. J Clin Invest 1993; 91:1214-1218.
41. Moshous D, Li L, de Chasseval R et al. A new gene involved in DNA double-strand break repair and V(D)J recombination is located on human chromosome 10p. Hum Mol Genet 2000; 9(4):583-588.
42. Nicolas N, Moshous D, Papadopoulo D et al. A human SCID condition with increased sensitivity to ionizing radiations and impaired V(D)J rearrangements defines a new DNA Recombination/Repair deficiency. J Exp Med 1998; 188:627-634.
43. Nicolas N, Finnie NJ, Cavazzana-Calvo M et al. Lack of detectable defect in DNA double-strand break repair and DNA-dependant protein kinase activity in radiosesitive human severe combined immunodeficiency fibroblasts. Eur J Immunol 1996; 26:1118-1122.
44. Li L, Drayna D, Hu D et al. The gene for severe combined immunodeficiency disease in Athabascan-speaking Native Americans is located on chromosome 10p. Am J Hum Genet 1998; 62(1):136-44.
45. Moshous D, Callebaut I, de Chasseval R et al. ARTEMIS, a Novel DNA Double-Strand Break Repair/V(D)J Recombination Protein, is Mutated in Human Severe Combined Immune Deficiency. Cell 2001; 105:177-186.
46. Rooney S, Sekiguchi J, Zhu C et al. Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. Mol Cell 2002; 10(6):1379-90.
47. Li L, Salido E, Zhou Y et al. Targeted disruption of the Artemis murine counterpart results in SCID and defective V(D)J recombination that is partially corrected with bone marrow transplantation. J Immunol 2005; 174(4):2420-8.
48. Gennery AR, Hodges E, Williams AP et al. Omenn's syndrome occurring in patients without mutations in recombination activating genes. Clin Immunol 2005; 116(3):246-56.
49. Evans PM, Woodbine L, Riballo E et al. Radiation-induced delayed cell death in a hypomorphic Artemis cell line. Hum Mol Genet 2006; 15(8):1303-11.
50. Aravind L. An evolutionary classification of the metallo-beta-lactamase fold proteins. In Silico Biol 1999; 1(2):69-91.
51. Ma Y, Pannicke U, Schwarz K et al. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell 2002; 108(6):781-94.
52. Lenain C, Bauwens S, Amiard S et al. The Apollo 5′ exonuclease functions together with TRF2 to protect telomeres from DNA repair. Curr Biol 2006; 16(13):1303-10.
53. van Overbeek M, de Lange T. Apollo, an Artemis-related nuclease, interacts with TRF2 and protects human telomeres in S phase. Curr Biol 2006; 16(13):1295-302.
54. Inagaki K, Ma C, Storm TA et al. A Role of DNA-PKcs and Artemis in Opening Viral DNA Hairpin Termini in Various Tissues in Mice. J Virol 2007; 81(20):11304-21
55. Callebaut I, Moshous D, Mornon JP et al. Metallo-β-lactamase fold within nucleic acids processing enzymes: the β-CASP family. Nucl Acid Res 2002; 30:3592-3601.
56. Poinsignon C, Moshous D, Callebaut I et al. The Metallo-β-Lactamase/β-CASP Domain of Artemis Constitutes the Catalytic Core Required for V(D)J Recombination. J Exp Med 2004; 199:315-321.

57. Pannicke U, Ma Y, Hopfner KP et al. Functional and biochemical dissection of the structure-specific nuclease ARTEMIS. EMBO J 2004; 23(9):1987-97.
58. Niewolik D, Pannicke U, Lu H et al. DNA-PKcs dependence of Artemis endonucleolytic activity, differences between hairpins and 5′ or 3′ overhangs. J Biol Chem 2006; 281(45):33900-9.
59. Ishikawa H, Nakagawa N, Kuramitsu S et al. Crystal structure of TTHA0252 from Thermus thermophilus HB8, a RNA degradation protein of the metallo-beta-lactamase superfamily. J Biochem (Tokyo) 2006; 140(4):535-42.
60. Mandel CR, Kaneko S, Zhang H et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. Nature 2006; 444(7121):953-6.
61. Poinsignon C, de Chasseval R, Soubeyrand S et al. Phosphorylation of Artemis following irradiation-induced DNA damage. Eur J Immunol 2004; 34(11):3146-55.
62. Riballo E, Kuhne M, Rief N et al. A pathway of double-strand break rejoining dependent upon ATM, Artemis and proteins locating to gamma-H2AX foci. Mol Cell 2004; 16(5):715-24.
63. Zhang X, Succi J, Feng Z et al. Artemis is a phosphorylation target of ATM and ATR and is involved in the G2/M DNA damage checkpoint response. Mol Cell Biol 2004; 24(20):9207-20.
64. Chen L, Morio T, Minegishi Y et al. Ataxia-telangiectasia-mutated dependent phosphorylation of Artemis in response to DNA damage. Cancer Sci 2005; 96(2):134-41.
65. Ma Y, Pannicke U, Lu H et al. The DNA-dependent protein kinase catalytic subunit phosphorylation sites in human Artemis. J Biol Chem 2005; 280(40):33839-46.
66. Wang J, Pluth JM, Cooper PK et al. Artemis deficiency confers a DNA double-strand break repair defect and Artemis phosphorylation status is altered by DNA damage and cell cycle progression. DNA Repair (Amst) 2005; 4(5):556-70.
67. Soubeyrand S, Pope L, De Chasseval R et al. Artemis phosphorylated by DNA-dependent protein kinase associates preferentially with discrete regions of chromatin. J Mol Biol 2006; 358(5):1200-11.
68. Goodarzi AA, Yu Y, Riballo E et al. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. EMBO J 2006; 25(16):3880-9.
69. Drouet J, Frit P, Delteil C et al. Interplay between Ku, Artemis and the DNA-dependent protein kinase catalytic subunit at DNA ends. J Biol Chem 2006; 281(38):27784-93.
70. Rogakou EP, Pilch DR, Orr AH et al. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. J Biol Chem 1998; 273(10):5858-68.
71. Darroudi F, Wiegant W, Meijers M et al. Role of Artemis in DSB repair and guarding chromosomal stability following exposure to ionizing radiation at different stages of cell cycle. Mutat Res 2007; 615(1-2):111-24.
72. Povirk LF, Zhou T, Zhou R et al. Processing of 3′-phosphoglycolate-terminated DNA double strand breaks by Artemis nuclease. J Biol Chem 2007; 282(6):3547-58.
73. Sancar A, Lindsey-Boltz LA, Unsal-Kacmaz K et al. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. Annu Rev Biochem 2004; 7339-85.
74. Geng L, Zhang X, Zheng S et al. Artemis links ATM to G2/M checkpoint recovery via regulation of Cdk1-cyclin B. Mol Cell Biol 2007; 27(7):2625-35.
75. Krempler A, Deckbar D, Jeggo PA et al. An imperfect G2M checkpoint contributes to chromosome instability following irradiation of S and G2 phase cells. Cell Cycle 2007; 6(14):1682-6.
76. Deckbar D, Birraux J, Krempler A et al. Chromosome breakage after G2 checkpoint release. J Cell Biol 2007; 176(6):749-55.
77. O'Driscoll M, Cerosaletti KM, Girard PM et al. DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. Mol Cell 2001; 8(6):1175-85.
78. Ben-Omran TI, Cerosaletti K, Concannon P et al. A patient with mutations in DNA Ligase IV: clinical features and overlap with Nijmegen breakage syndrome. Am J Med Genet A 2005; 137(3):283-7.
79. Buck D, Moshous D, de Chasseval R et al. Severe combined immunodeficiency and microcephaly in siblings with hypomorphic mutations in DNA ligase IV. Eur J Immunol 2006; 36(1):224-35.
80. Enders A, Fisch P, Schwarz K et al. A severe form of human combined immunodeficiency due to mutations in DNA ligase IV. J Immunol 2006; 176(8):5060-8.
81. van der Burg M, van Veelen LR, Verkaik NS et al. A new type of radiosensitive T-B-NK+ severe combined immunodeficiency caused by a LIG4 mutation. J Clin Invest 2006; 116(1):137-45.
82. Buck D, Malivert L, de Chasseval R et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 2006; 124(2):287-99.
83. Revy P, Buck D, le Deist F et al. The repair of DNA damages/modifications during the maturation of the immune system: lessons from human primary immunodeficiency disorders and animal models. Adv Immunol 2005; 87237-95.
84. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 2006; 124(2):301-13.

85. Zha S, Alt FW, Cheng HL et al. Defective DNA repair and increased genomic instability in Cernunnos-XLF-deficient murine ES cells. Proc Natl Acad Sci USA 2007; 104(11):4518-23.
86. Dai Y, Kysela B, Hanakahi LA et al. Nonhomologous end joining and V(D)J recombination require an additional factor. Proc Natl Acad Sci USA 2003; 100(5):2462-7.
87. Hentges P, Ahnesorg P, Pitcher RS et al. Evolutionary and functional conservation of the DNA non-homologous end-joining protein, XLF/Cernunnos. J Biol Chem 2006; 281(49):37517-26.
88. Sibanda BL, Critchlow SE, Begun J et al. Crystal structure of an Xrcc4-DNA ligase IV complex. Nat Struct Biol 2001; 8(12):1015-9.
89. Junop MS, Modesti M, Guarne A et al. Crystal structure of the Xrcc4 DNA repair protein and implications for end joining. EMBO J 2000; 19(22):5962-70.
90. Callebaut I, Malivert L, Fischer A et al. Cernunnos interacts with the XRCC4 x DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. J Biol Chem 2006; 281(20):13857-60.
91. Lu H, Pannicke U, Schwarz K et al. Length-dependent binding of human XLF to DNA and stimulation of XRCC4.DNA ligase IV activity. J Biol Chem 2007; 282(15):11155-62.
92. Deshpande RA, Wilson TE. Modes of interaction among yeast Nej1, Lif1 and Dnl4 proteins and comparison to human XLF, XRCC4 and Lig4. DNA Repair (Amst) 2007; 6(10):1507-16.
93. Frank-Vaillant M, Marcand S. NHEJ regulation by mating type is exercised through a novel protein, Lif2p, essential to the ligase IV pathway. Genes Dev 2001; 15(22):3005-12.
94. Kegel A, Sjostrand JO, Astrom SU. Nej1p, a cell type-specific regulator of nonhomologous end joining in yeast. Curr Biol 2001; 11(20):1611-7.
95. Valencia M, Bentele M, Vaze MB et al. NEJ1 controls nonhomologous end joining in Saccharomyces cerevisiae. Nature 2001; 414(6864):666-9.
96. Cavero S, Chahwan C, Russell P. Xlf1 is required for DNA repair by nonhomologous end joining in Schizosaccharomyces pombe. Genetics 2007; 175(2):963-7.
97. Wu PY, Frit P, Malivert L et al. Interplay between Cernunnos/XLF and nonhomologous end-joining proteins at DNA ends in the cell. J Biol Chem 2007; 282(44):31937-43.
98. Ahnesorg P, Jackson SP. The nonhomologous end-joining protein Nej1p is a target of the DNA damage checkpoint. DNA Repair (Amst) 2007; 6(2):190-201.
99. Grawunder U, Zimmer D, Leiber MR. DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains. Curr Biol 1998; 8(15):873-6.
100. Teo SH, Jackson SP. Lif1p targets the DNA ligase Lig4p to sites of DNA double-strand breaks. Curr Biol 2000; 10(3):165-8.
101. Grawunder U, Wilm M, Wu X et al. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. Nature 1997; 388(6641):492-5.
102. Tsai CJ, Kim SA, Chu G. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. Proc Natl Acad Sci USA 2007; 104(19):7851-6.
103. Gu J, Lu H, Tsai AG et al. Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. Nucleic Acids Res 2007; 35(17):5755-62.
104. Gao Y, Sun Y, Frank KM et al. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. Cell 1998; 95(7):891-902.
105. Frank KM, Sekiguchi JM, Seidl KJ et al. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. Nature 1998; 396(6707):173-7.
106. Franco S, Alt FW, Manis JP. Pathways that suppress programmed DNA breaks from progressing to chromosomal breaks and translocations. DNA Repair (Amst) 2006; 5(9-10):1030-41.
107. Moshous D, Pannetier C, Chasseval Rd R et al. Partial T- and B-lymphocyte immunodeficiency and predisposition to lymphoma in patients with hypomorphic mutations in Artemis. J Clin Invest 2003; 111(3):381-7.

CHAPTER 5

Large-Scale Chromatin Remodeling at the Immunoglobulin Heavy Chain Locus:

A Paradigm for Multigene Regulation

Daniel J. Bolland, Andrew L. Wood and Anne E. Corcoran*

Abstract

V(D)J recombination in lymphocytes is the cutting and pasting together of antigen receptor genes in cis to generate the enormous variety of coding sequences required to produce diverse antigen receptor proteins. It is the key role of the adaptive immune response, which must potentially combat millions of different foreign antigens. Most antigen receptor loci have evolved to be extremely large and contain multiple individual V, D and J genes. The immunoglobulin heavy chain (*Igh*) and immunoglobulin kappa light chain (*Igk*) loci are the largest multigene loci in the mammalian genome and V(D)J recombination is one of the most complicated genetic processes in the nucleus. The challenge for the appropriate lymphocyte is one of macro-management—to make all of the antigen receptor genes in a particular locus available for recombination at the appropriate developmental time-point. Conversely, these large loci must be kept closed in lymphocytes in which they do not normally recombine, to guard against genomic instability generated by the DNA double strand breaks inherent to the V(D)J recombination process. To manage all of these demanding criteria, V(D)J recombination is regulated at numerous levels. It is restricted to lymphocytes since the Rag genes which control the DNA double-strand break step of recombination are only expressed in these cells. Within the lymphocyte lineage, immunoglobulin recombination is restricted to B-lymphocytes and TCR recombination to T-lymphocytes by regulation of locus accessibility, which occurs at multiple levels. Accessibility of recombination signal sequences (RSSs) flanking individual V, D and J genes at the nucleosomal level is the key micro-management mechanism, which is discussed in greater detail in other chapters. This chapter will explore how the antigen receptor loci are regulated as a whole, focussing on the *Igh* locus as a paradigm for the mechanisms involved. Numerous recent studies have begun to unravel the complex and complementary processes involved in this large-scale locus organisation. We will examine the structure of the *Igh* locus and the large-scale and higher-order chromatin remodelling processes associated with V(D)J recombination, at the level of the locus itself, its conformational changes and its dynamic localisation within the nucleus.

*Corresponding Author: Anne E. Corcoran—Laboratory of Chromatin and Gene Expression, Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK.
Email: anne.corcoran@bbsrc.ac.uk

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

Introduction

B-Cell Development

In order to generate the primary repertoire of immunoglobulins and T-cell receptors, antigen receptor loci undergo variable, diversity and joining (V(D)J) recombination in B- and T-lymphocytes. This involves generation of DNA double strand breaks at recombination signal sequences (RSSs) flanking individual genes, followed by removal of the intervening DNA and juxtapositioning and ligation of the recombining gene segments. This process is regulated at several levels. First, recombination is catalyzed by a recombinase complex containing the protein products of the recombinase activating genes *Rag1* and *Rag2*.[1] RAG expression is restricted to precursor lymphocytes, thereby restricting V(D)J recombination to these cells. Second, within precursor lymphocytes, this process is strictly lineage-specific with heavy (*Igh*) and light (*Igk* and *Igl*) immunoglobulin loci only fully recombining in B-lymphocytes and T-cell receptor loci (*Tcra*, *Tcrb*, *Tcrg* and *Tcrd*) only recombining in T-cells. Third, within lineages, loci are recombined in a precise order. Recombination of the *Igh* locus in pro-B-cells is the earliest step in the generation of the mature antibody repertoire in B-lymphocytes and is followed by *Igk* and then *Igl* recombination in preB-cells. Fourth, the order is also strictly maintained within loci: D_H-to-J_H recombination occurs on both *Igh* alleles before V_H-to-DJ_H recombination takes place.[2] Finally, RAG activity is targeted to RSSs flanking individual V, D and J genes. Apart from restriction to lymphocytes by restricted RAG expression, this ordered regulation is effected by several levels of immunoglobulin locus accessibility.

Description of the Igh *Locus*

This need for multiple levels of regulation is both necessitated and complicated by the enormous size of the antigen receptor loci. The mouse *Igh* and *Igk* loci are the largest multigene loci known, with sizes of 3Mb and 3.2 Mb respectively.[3,4] The *Igh* locus of the C57BL/6 mouse has recently been completely assembled and annotated. It comprises 195 V_H genes spanning 2.5Mb, 10 D_H genes (~60kb), 4 J_H genes (2kb) and 8 constant (C_H) genes (200kb) (Fig. 1).[3,5] The V genes are organized into 16 families of varying sizes, based on sequence homology. The majority are functional, but a large proportion (85) are classed as pseudogenes, some of which nevertheless recombine, although they do not make functional Ig polypeptides. All of the functional V, D and J genes are used in multiple different combinations and this large choice of V, D and J recombination partners provides the first step in immunoglobulin diversity. However there is a bias in recombination frequency between the 3' and 5' ends of the V region i.e., the 3' end is recombined more frequently in fetal liver and in the earliest bone marrow B-cells. The extent of the bias varies between mouse strains and recombination frequency normalises in later B-cells.[6-8] Large-scale mechanisms which may contribute to the bias will be discussed below.

Each V and D gene has its own promoter and all genes are transcribed in the same orientation (Fig. 1), although this is not the case for all antigen receptor loci. Promoters have several features in common, but also family-specific differences which may be a factor in observed family-specific differences in recombination frequency.[3] The human Igh locus is smaller (1Mb) and contains only 123 V genes, 79 of which are pseudogenes.[9] The V regions of the *Igh* and other antigen receptor loci are believed to have evolved from much smaller V gene clusters that were frequently duplicated, possibly due to ability of the Rag enzymes to act as general transposases.[10,11] Consequently even within species there are significant differences in numbers and family distribution of V genes, particularly in the mouse.[3,12] For example the 7183 gene family at the 3' end of the V region has 21 V genes in the C57BL/6 strain and 49 V genes in the 129 strain. This is an extremely important consideration when comparing *Igh* locus recombination between mouse strains. In the future it is likely that studies on the C57BL/6 strain will predominate as this is the strain in which the mouse genome was sequenced and thus contains all other relevant sequence information.

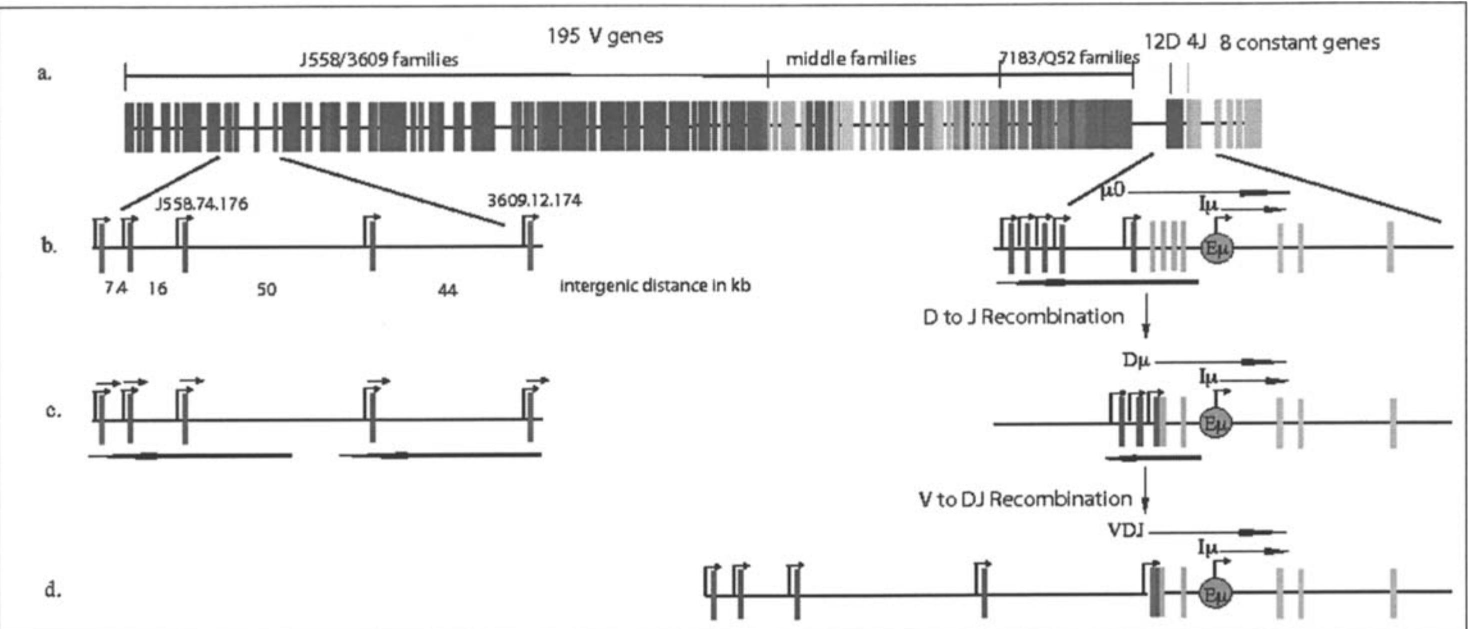


Figure 1. Schematic drawing of the mouse C57BL/6 *Igh* locus and noncoding RNA transcription during V(D)J recombination. a) Approximate scale drawing of the V, D and J gene families, with different colors/shades representing different V gene families. Approximate distances between genes are depicted, albeit for clarity the genes occupy a disproportionately large linear space. For example the J558/3609 region which occupies the 5′ 1.5Mb contains widely spaced J558 genes interspersed with 3609 genes in the 5′ 1Mb (see zoomed out picture), but highly clustered J558 genes with no intervening 3609 genes in the 3′ 500kb. Transcripts (sense above, antisense below the genes) are presented by horizontal arrows. b) Magnified picture of the germline *Igh* locus, showing noncoding RNA transcripts. Numbers above the V genes indicate gene identity (see ref. 3). c) Magnified picture of the DJ recombined locus, as above. d) Magnified picture of the VDJ recombined locus, as above.

Chromatin Remodeling

How is such an enormous piece of DNA manipulated in the nucleus to ensure that its many genes are accessible for V(D)J recombination in pro-B-cells, but not in T-cells or later stage B-cells? While recombination itself is a genetic process i.e., alterations are made in the DNA sequence of the locus, it is regulated by a multitude of epigenetic processes i.e., heritable changes in chromatin structure that do not involve a change to the primary sequence. It is important to bear in mind that structurally this extremely long DNA sequence is not simply a piece of string, but occupies a 3-dimensional space in the nucleus. It is estimated that the linear length of DNA helix contained in a mammalian genome is 3 meters and this must be accommodated in a cell nucleus with a diameter of 5-10μm. This is achieved at the basic level by wrapping the DNA helix around the histone octamer in the nucleosome, followed by several levels of higher order folding of nucleosomes over each other, in ways that are not well understood (Fig. 2). To facilitate general gene transcription, this higher order chromatin must first be unravelled to achieve a more open and ultimately single nucleosomal structure. This kind of multi-tiered regulation also controls V(D)J recombination[13] and recent studies have explored the extent to which these mechanisms are involved in V(D)J recombination. This chapter will explore several aspects in detail—noncoding RNA transcription, nuclear localization and regulatory elements, while placing these in context with other processes including histone modification, which will be explored in detail in other chapters in this volume. We will explore what is currently known, what current studies may predict and what the future directions are likely to be.

Noncoding RNA Transcription

Contrary to the 'central dogma' that 'DNA makes RNA makes protein', numerous genome-wide transcriptional analyses have estimated that over half of all transcribed mammalian genomic sequences are nonprotein-coding[14] and some of this transcription is predicted to play key roles in gene regulation. Notably, over 20 years ago, the *Igh* locus was the one of the first loci shown to express noncoding RNAs. This transcription was originally termed 'sterile' or 'germline' to distinguish it from coding transcription from V(D)J recombined genes. In the *Igh* locus, the first germline transcripts occur before D_H-to-J_H recombination and initiate from two regions; the intronic enhancer E_μ (I_μ transcript)[15] and from a promoter, PDQ52, immediately upstream of the most 3' D_H gene segment, DQ52 ($_\mu 0$ transcript)[16] (Fig. 1). Following D_H-to-J_H recombination, the DJ_H gene segment produces D_μ transcripts[17] and sense germline transcription initiates over the V_H genes (Fig. 1).[18,19] Subsequently, noncoding RNA transcripts have been identified in all antigen receptor loci across gene segments competent for recombination.[20] The discovery of V_H gene germline transcription formed the basis of the accessibility hypothesis, which proposed that

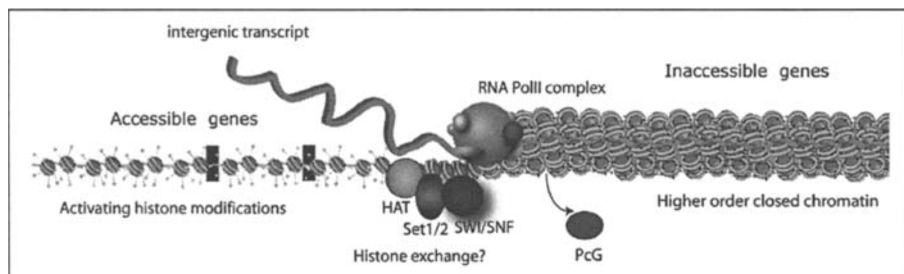


Figure 2. Model of intergenic transcription. The RNA Pol II complex, depicted as a sphere, with associated smaller spheres denoting transcription factors of the basal complex, is pictured processing through closed chromatin, recruiting activating chromatin remodelling factors and promoting egress of PcG (Polycomb) proteins. HAT: histone acetyltransferase; Set1/2: members of Trithorax family of histone H3 HMTs; SWI/SNF: SWItch/sucrose nonfermentable.

lineage- and stage-specificity of recombination are regulated by differential chromatin accessibility of antigen receptor gene segments to the recombinase machinery, with germline transcription associated with open chromatin.[18,21] However, a function for V_H germline transcription has not been formally demonstrated and it has been argued that it may be a secondary effect of the V_H gene promoters becoming accessible for V_H-to-DJ_H recombination. Neither have functions yet been assigned to the $_{\mu}0$ and I_{μ} transcripts. However, quantitative RNA-FISH visualization of I_{μ} transcription[22,23] have classed this transcript as a 'supergene' i.e., a gene that is transcribed almost continuously from both alleles in an individual nucleus.[24] This property applies to surprisingly few genes, β-globin among them. I_{μ} is the first noncoding 'supergene' to be identified and would more correctly be termed a 'super-transcription unit', since it is a noncoding, intergenic transcript. The possible implications of this high level transcription for the role of noncoding RNA transcription in the *Igh* locus will be discussed below.

Intergenic Transcription

Recent studies suggest that intergenic transcription may play a role in opening up the *Igh* locus. In the large V region, the relatively small V genes (500bp) are separated by enormous intergenic distances (10-20kb).[3] The chromatin remodeling processes previously discovered are largely confined to V genes (germline transcription above, histone modifications, discussed in detail in other chapters). Such focused alterations are unlikely to be sufficient to open the closed chromatin conformation of the locus, the default state in nonB-cells[25] and additional large-scale processes were investigated. In numerous loci, including β-globin and the MHC complex, intergenic transcription delineates domains of modified chromatin that surround active genes and their regulatory elements.[26-29] RNA polymerase II (PolII) recruits a wide range of chromatin remodeling and histone-modifying factors, including histone acetyltransferases (HATs) and Set1 and Set2 histone methyltransferases (HMTs), required for histone modifications associated with gene activation.[30-34] Furthermore, transcription triggers histone turnover and the deposition of variant histone H3.3, enriched with active modifications.[35] Collectively these activities suggest several mechanisms by which the processing activity of elongating PolII complex can achieve chromatin accessibility.[36,37] Accordingly, intergenic transcription has been proposed to drive through repressive chromatin in several multigene loci, recruiting remodeling factors and opening up large chromatin domains into a poised state, thus facilitating further focused chromatin opening over genes to regulate gene expression (Fig. 2).[38] In several large developmentally regulated loci,[28] this is believed to occur by transcription-dependent[39,40] higher order chromatin remodeling and looping out of their chromosome territories.[41,42]

In many cases, intergenic transcription may only need to drive through once or twice to open up the chromatin. However, in other instances, including the Drosophila homeotic bithorax complex, continuous intergenic transcription is required to prevent binding of repressive Polycomb proteins with H3K27 HMT activity and to recruit activating Trithorax H3K4 HMTs.[43] Furthermore, many enhancers and Locus Control Regions undergo transcription, which is essential for activation of their target genes and thus transcription from an intergenic regulatory region can influence expression of a distal gene.[44]

Intergenic Transcription in the Mouse *Igh* Locus V Region

Analysis of transcription from genes and intergenic regions throughout the *Igh* V_H region, using RT-PCR to measure steady-state levels and RNA-FISH to visualize primary transcripts on individual alleles in single cells revealed that intergenic transcription occurs throughout the Igh V region. It is absent on germline alleles that have not yet recombined D_H to J_H in early B-cells, is expressed on the majority of D_HJ_H recombined alleles and disappears once V to DJ recombination has occurred. This tightly developmentally regulated pattern of expression is characteristic of a large-scale functional process. Furthermore, patterns of transcripts detected by RNA-FISH were extended over large regions, suggesting extensive transcription on individual alleles (Fig. 3).

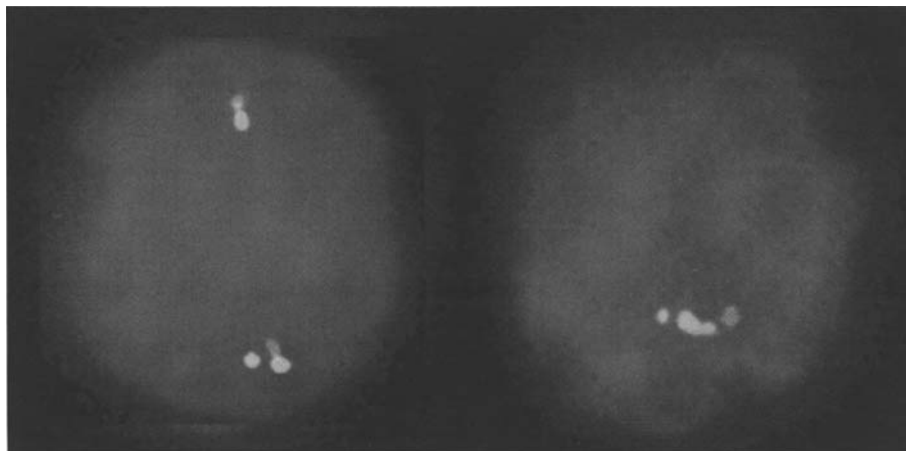


Figure 3. Visualization of antisense transcription in the *Igh* locus by RNA-FISH. Nuclei from ex vivo wild-type bone marrow sorted for Fraction B-cells (the majority of cells are DJ recombined). I_μ sense transcripts, hybridized with a single-stranded antisense probe, are detected by Texas Red (red/light grey punctate signals). J558 gene family antisense transcripts, hybridized with a single-stranded sense probe, are detected by fluoroisothiocyanate (FITC) (green/dark grey extended signals). Nuclei are counterstained with 4', 6-diamidino-2-phenylindole (blue). (adapted with permission from ref. 22). A color version of this figure is available at www.landesbioscience.com/curie.

Antisense Transcription

Surprisingly intergenic transcription only occurred on the antisense strand, where transcription also occurred over the V_H genes. Antisense transcription has classically been associated with transcription repression in imprinted loci, in which it appears to silence gene expression in *cis* from the allele on which it is expressed. The best studied example is Air (Antisense to Igf2 receptor) transcription (108kb transcript), which silences expression of the Igf2R, with which it partially overlaps, in *cis*.[45] Antisense transcription has also been documented in several lower eukaryotic systems to generate dsRNA and heterochromatin formation.[46,47] However, it is now thought that the majority of mammalian transcription units display overlapping sense and antisense transcription.[48] This high incidence and co-ordinate regulation of many sense-antisense pairs, indicates that antisense transcription is involved in mechanisms other than its classical association with transcriptional repression.[45,49] For example, antisense transcription across the yeast PHO5 gene promoter is required to increase the rate of transcription and is believed to evict histones to enable greater access of RNAPolII to the gene.[50] In the mammalian HOXA cluster, antisense intergenic transcription is required to activate neighboring HOX genes, in part by disrupting interaction with repressive PcG complexes.[51] These examples may be the first of many in which antisense transcription plays an activating role.

Antisense Transcription in the *Igh* Locus V Region

In the *Igh* locus, the absence of V region antisense transcription on germline alleles argues against this transcription keeping the V_H region closed, since it would have to be present before D_H to J_H recombination to do this. Rather, it it is consistent with a role for intergenic transcription in opening up the V_H region and thus it doesn't appear on germline alleles, since the V_H region must be kept closed until D_HJ_H recombination has taken place. Furthermore antisense transcription is biallelic, arguing against a monoallelic mechanism of silencing one allele to prevent recombination. The expression pattern of antisense transcription in the *Igh* locus thus argues in favor of its having an activating rather than a repressive role in V(D)J recombination. Further, this transcription is not controlled by V_H gene promoters and thus cannot be regarded as a by-product of the activation of

these promoters for V(D)J recombination. This is the first evidence in support of a functional role for germline transcription in *Igh* V(D)J recombination. We proposed this large-scale transcription remodels the V_H region to facilitate accessibility for V_H-to-DJ_H recombination, perhaps by directing chromatin remodeling factors to direct other changes in chromatin structure that precede V(D)J recombination (Fig. 4).[22] These occur mostly over the V_H genes and include loss of histone H3K9

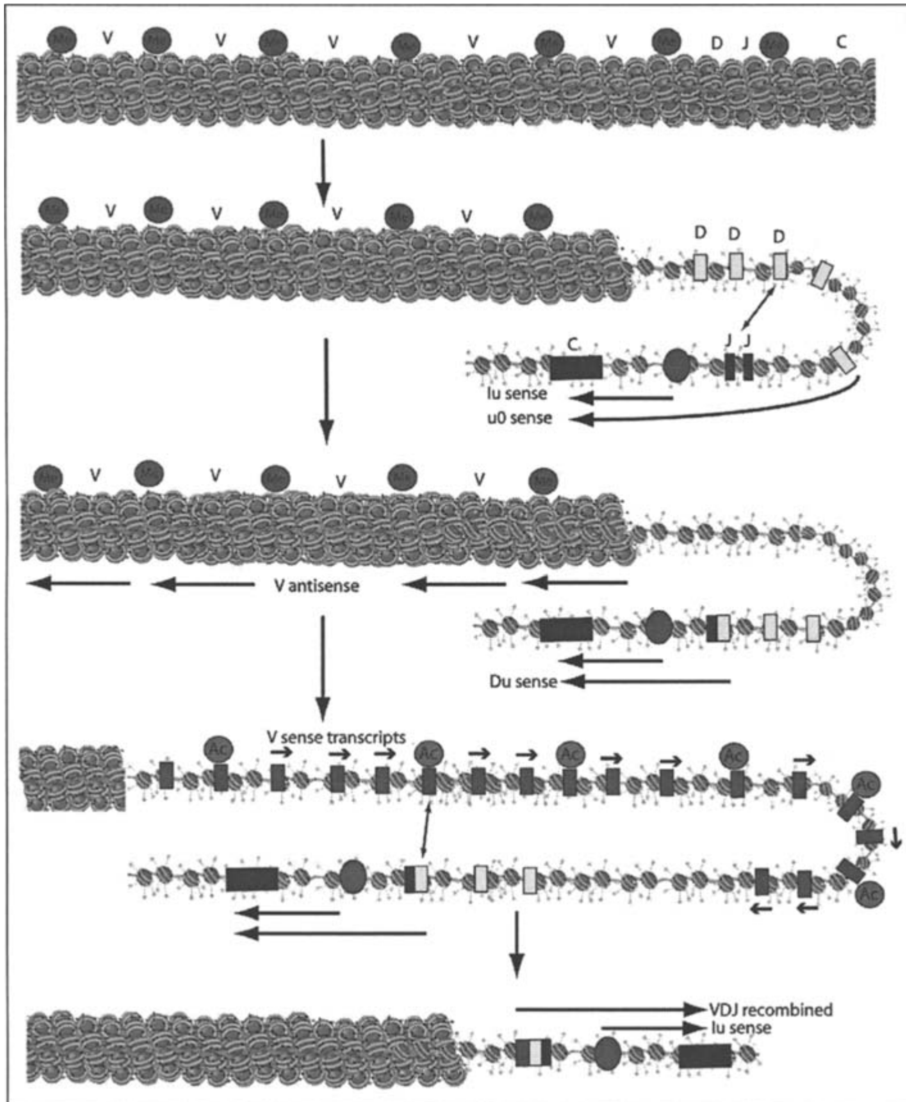


Figure 4. Model of role of antisense intergenic transcription in *Igh* V(D)J recombination. Schematic of order of events, depicting alterations in chromatin structure. Key: Multiple red/light grey boxes: V genes; yellow/light grey boxe: D genes; blue/dark grey boxes: J genes; E_μ: green oval; large rectangular box: constant region; black arrows: sense/antisense transcripts; Me in red circle: repressive histone modifications; Ac in green circle: activating histone modifications. A color version of this figure is available online at www.landesbioscience.com/curie.

methylation, acetylation of histones H3 and H4, markers of accessible chromatin, histone H3.3 exchange and methylation of H3 lysine 27 (H3-27).[25,52-55]

Antisense and Intergenic Transcription in the *Igh* D Region

The discovery of intergenic antisense transcription over the *Igh* V region before V_H-to-D_HJ_H recombination raised the question of whether similar transcriptional processes precede other V(D)J recombination events. Antisense intergenic transcription also occurs throughout the D_H (60kb) and J_H regions of the mouse *Igh* locus in pro-B-cells poised for D_H-to-J_H recombination and is thus a widespread process during V(D)J recombination.[56] It is activated on germline alleles before D_H-to-J_H recombination. Notably, it initiates near to and is regulated by the intronic enhancer E_μ.[56] E_μ was originally proposed to regulate V_H to DJ_H recombination.[57,58] However, recent studies have shown that targeted deletion of E_μ causes a defect in D_H-to-J_H recombination, suggesting that E_μ primarily regulates this process and that defects in V_H to DJ_H recombination may be secondary to this earlier defect.[59,60] It is not yet understood how E_μ regulates D_H to J_H recombination. Transcription of the I_μ 'supergene' initiates immediately downstream. Deletion of E_μ results in loss of both I_μ sense[60] and D region antisense transcription, up to 50 kb away.[56] This suggests that E_μ controls D_H-to-J_H recombination at least in part by activating germline Igh transcription and that in particular, the processivity of the antisense transcription renders the D_H and J_H regions accessible for D_H-to-J_H recombination (Fig. 4).

This model is supported by concomitant increases in DNase I sensitivity, histone H3 and H4 acetylation, H3K4 methylation and nucleosome remodeling enzymes over D_H and J_H genes in pro-B-cells.[52,61,62] Histone acetylation is widespread throughout the D_H region,[52] but is highest over the J_H region and the DQ52 gene,[52,61] which is preferentially used in early D_H-to-J_H recombination.[63] This model is in agreement with a recent suggestion that the region encompassing DQ52, the four J genes and E_μ forms a separate chromatin domain to the rest of the D_H region.[62] Strikingly, DQ52 is the only D_H gene that expresses both sense and antisense germline transcripts and this transcription overlap extends into the J_H region.[56] These data suggest strongly that the transcripts do not produce dsRNAs that lead to heterochromatin. Indeed, they are coordinately up-regulated by E_μ. Additionally, there is no sense germline transcription in the remainder of the D_H region, precluding dsRNA formation.[56,64] Nevertheless, a recent report of active retention of repressive histone marks over the middle D_H genes has led to the opposite hypothesis that antisense transcription may contribute to repression of these genes, by formation of dsRNA and Dicer-mediated heterochromatinization, albeit no dsRNA was detected.[64] Definitive resolution of these opposing models must await clarification of the functional role of antisense transcription by targeted removal of this transcription in vivo. Similar gene targeting studies have shown that intergenic transcription is functionally required for V(D)J recombination at the *Tcra* locus, but in this case it originates from the sense strand.[65] This suggests that the strand origin is not important, which supports the model that the processing activity is the key function of this transcription.

D_H antisense transcripts initiate on germline alleles and V_H transcripts on DJ recombined alleles and D_H and V_H antisense transcripts are rarely associated on individual alleles.[56] Thus there is a stepwise progression of antisense intergenic transcription, in a strikingly similar pattern to the stepwise progression of active histone modifications during *Igh* V(D)J recombination. These occur first over the D_HJ_H region, then sequentially over the 3' end, the middle region and the 5' end of the V_H region.[53,66-68] Thus antisense intergenic transcription may facilitate the exchange of repressive histone marks associated with the locus in nonB-cells with active histone marks, perhaps by histone exchange in favor of active histones e.g., H3.3 (Fig. 4).[25] Notably in the *Tcra* locus, intergenic transcription has been shown to increase active histone marks over genes.[65]

Subnuclear Relocalisation

In addition to these localized and large-scale epigenetic changes over the *Igh* locus, the location of the locus in the nucleus has an enormous impact on its recombination potential. In nonB-lymphoid cells, the *Igh* and *Igk* loci are maintained at the nuclear periphery, generally

regarded as a repressive chromatin environment, although it is not clear whether the *Igh* is specifically associated with repressive chromatin at this location.[69] The D_HJ_H distal J558 V_H genes are oriented towards the nuclear envelope and the locus is effectively 'tethered' at the periphery via the J558 genes, while the D_HJ_H region is oriented towards the centre of the nucleus, which may contribute to D_HJ_H occurring before V_H to D_HJ_H recombination.[70] In early B-cells undergoing V(D)J recombination, both *Igh* and *Igk* alleles are repositioned to the euchromatic interior of the nucleus, a region permissive for transcription.[69] The relocation is dependent on interleukin-7 receptor signalling, but is independent of RAG[69] or Pax5[71] expression. This nuclear repositioning appears to be sufficient for D_HJ_H recombination and V_H to D_HJ_H recombination of D_H-proximal V_H genes in the *Igh* locus.

3-Dimensional Alterations in Chromatin Structure

However, to achieve recombination of middle and D_H-distal V_H genes, central nuclear repositioning is not sufficient, presumably due to the enormous size of the locus. An additional process, termed locus contraction, is required. This juxtaposes the distal V_H genes with the D_HJ_H recombined gene segment in pro-B-cells and is mediated by higher-order chromatin looping of individual IgH subdomains.[72,73] It is regulated by the transcription factor Pax5 (Fig. 5).[71] Pax5 is the pivotal transcription factor that regulates establishment and maintenance of B-lymphocyte identity and its absence prevents recombination of middle and D-distal genes.[74] Looping is also regulated by the multifunctional transcription factor, YY1, which binds E_μ.[75] It is unclear how

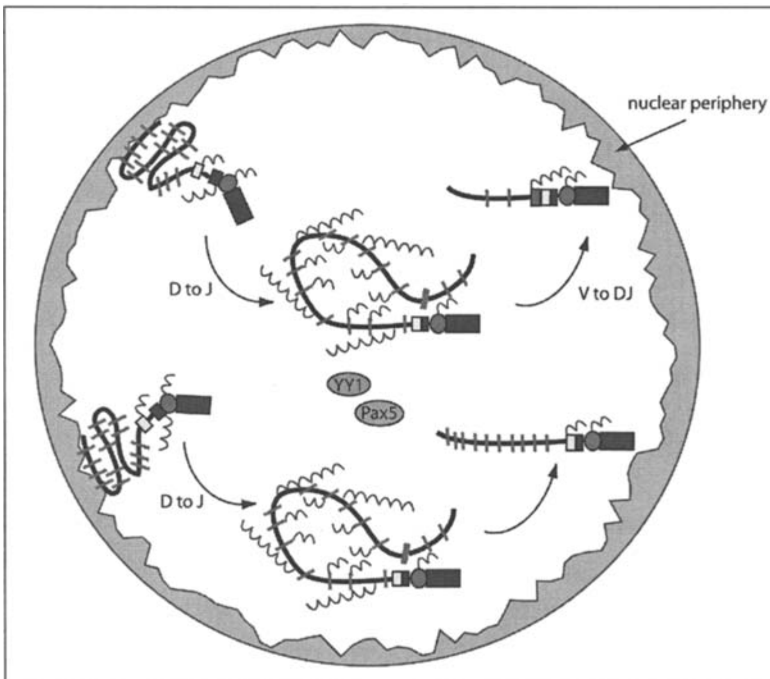


Figure 5. Nuclear organisation of the *Igh* locus. The sequential stages of *Igh* V(D)J recombination are represented in the context of the spatial location of the *Igh* loci in the nucleus and their large-scale conformation changes. The locus is initially tethered at the nuclear periphery via the 5′ end of the V region. Key: Multiple red/light grey lines: V genes; yellow/light grey box: D region; blue/dark grey box: J region; E_μ: green oval; large rectangular box: constant region; short (blue) squiggles: sense transcripts; long (purple) squiggles: antisense transcripts. A color version of this figure is available online at www.landesbioscience.com/curie.

either Pax5 or YY1 enable DNA looping. Neither is required for germline transcription or histone acetylation of these genes,[67] suggesting that these processes are either necessary but not sufficient, or independent of looping. One possibility is that YY1 may recruit other parts of the locus to the enhancer and future studies on the role of E_μ in looping will be informative. YY1 binds Ezh2, a polycomb group protein with H3K27 HMTase activity, although this binding has not yet been shown in B-cells. Ezh2 is required also for recombination of distal V genes.[55] Its mechanism of action is currently unclear, but intriguingly it also appears to be required for DNA looping of the Igh V_H region (A Tarahkovsky, personal communication).

Transcription Factories

A large body of recent evidence has shown that transcription does not occur homogeneously throughout the nucleus, but appears to be concentrated in sub-nuclear foci of active RNA PolII complexes, termed transcription factories.[76,77] Individual transcription factories are believed to contain up to ten RNA PolII complexes and to transcribe several genes simultaneously.[76] These genes can be up to 40 MB apart on the same chromosome and even on separate chromosomes.[23,78] These are dynamic interactions that reflect the frequency of transcription of individual genes.[78] Most genes are not transcribed continuously, but rather switched on and off stochastically.[79] The I_μ 'supergene' is transcribed almost all of the time in both proB and mature B-cells and is thus almost continuously associated with a transcription factory.[22,23] It has recently been shown that enhancers can relocate genes away from the nuclear periphery by recruiting them to a transcription factory.[80] In a similar manner, E_μ may promote nuclear relocation by recruiting the DJ region to a transcription factory in the nuclear interior, where E_μ facilitated transcription may then keep the D_HJ_H region in the transcription factory, providing a relatively stable focal point for DNA looping.

Biased Recombination Frequency Explained by Numerous Mechanisms

The studies above provide several large-scale contributing reasons for preferential recombination of 3' V_H genes in early B-cells. First, the *Igh* is tethered at the nuclear periphery in nonB-cells via the J558 genes at the 5' end, thus the 3' genes are oriented towards and relocated into the central euchromatin first. Following relocation, it appears that proximal V_H genes are less dependent on DNA looping of the V_H region for recombination, presumably due to their proximity to the D_HJ_H region.[71,72] Furthermore, all the factors that regulate looping (Pax5, YY1, Ezh2) are only required for recombination of distal V_H genes. The interleukin 7 receptor is also required for recombination of 5' genes, but not 3' V_H genes in the bone marrow.[19] Since it activates germline transcription over 5' V_H genes, but not 3' V_H genes, it was proposed that it increased V region chromatin accessibility to the recombinase.[19] Subsequent studies have identified other contributory mechanisms regulated by the IL7R. It is required for relocation from the nuclear periphery[69] and histone acetylation of 5' V_H genes.[66,67]

Allelic Choice and Allelic Exclusion

Ultimately the goal of the B-lymphocyte is to express a $V_HD_HJ_H$ recombined *Igh* gene from only one allele at the cell surface. Surface expression of the immunoglobulin polypeptide is believed to lead to a feedback signaling cascade that silences the second allele, a mechanism termed allelic exclusion.[2] This ensures that each lymphocyte produces monoclonal antibodies that recognize a single antigen with high specificity. Several processes contribute to this monoallelic expression. In the *Igh* locus, V_H to D_HJ_H recombination is asynchronous—i.e., one allele undergoes recombination first. This reduces the danger of simultaneously producing two productive recombination events. However, unlike the *Igk* (see below), it is unclear how this allelic choice is achieved in the *Igh* locus. Relocation and antisense intergenic transcription appear to be biallelic. It is currently unclear whether locus contraction is mono or biallellic and further studies are required to reveal whether it plays a role in allelic choice.[72,73] However, it is clear that

the opening mechanisms required for V(D)J recombination are mirrored by a set of opposing processes designed to stop further V(D)J recombination once a productive recombination event has yielded a protein product. Histone acetylation is reduced over V_H genes,[54,81] sense and antisense germline transcription is lost,[22] locus de-contraction occurs.[72] These processes occur on both alleles. An additional mechanism occurs specifically on the second allele that has either yielded a nonproductive $V_HD_HJ_H$ rearrangement or has not yet managed to rearrange the V_H gene (D_HJ_H rearranged allele). In either case, the allele is believed to be recruited to repressive pericentromeric heterochromatin, which may preclude further V to DJ recombination.[72,82] It is recruited via the 5' end of the V region and silencing of the locus is not complete. The I_μ,[22] DJ rearranged[83] and sense germline transcripts from 3' V genes[84] continues to be transcribed. This is presumably because D to J recombination has already occurred on both alleles and thus only the V region needs to be prevented from further recombination.

Other Antigen Receptor Loci

We have focused on the *Igh* locus, which has proved to be a useful paradigm for other antigen receptor loci, since, with some exceptions, processes discovered in the *Igh* locus, also occur in other antigen receptor loci. For example, noncoding sense RNA transcription over V genes has been observed in most other antigen receptor loci.[20] Similarly relocation from the nuclear periphery and locus contraction by DNA looping has been reported in the *Igk*, *Tcra* and *Tcrb* loci.[69,72,85] However, it is not yet known how widespread the process of antisense and/or intergenic transcription is. The biggest difference between recombination of *Igh* and *Igk* is the order and nature of the events that ensure monoallelic expression. In contrast to the *Igh*, in which this appears to be controlled after V(D)J recombination, the *Igk* loci undergo several monoallelic processes before V to J recombination, which render one allele preferentially more available for the initial recombination event. One allele is preferentially DNA demethylated[86] and acquires active histone marks before V to J recombination, while the second allele remains DNA methylated and is recruited to heterochromatin before V to J recombination.[87]

Future Directions

Further studies are required to unequivocally determine the function of antisense intergenic transcription in the *Igh* locus in vivo. Furthermore is it the processivity of the transcription that is important, its strand-specificity, or indeed the transcripts themselves? These are also important considerations for other antigen receptor loci.

There is also little known about other chromatin remodeling processes in *Igh* intergenic regions. It is unclear whether noncoding RNA transcription is regulated by the same histone modifications as coding transcription. Since there are now more than 150 known histone modifications,[88] it will be important to explore the possibility that recombination may have a unique histone code which does not correspond to the code for transcription.

There is as yet no regulatory element defined for the *Igh* V region. However, a novel pro-B-cell specific HS site has recently been identified 5' of the V region.[89] It will be interesting to see if this element regulates V to DJ recombination, albeit initial characterization indicates a repressive role. How might this or another regulatory element function? It might activate V region antisense transcription or enable DNA looping by interacting with elements close to the DJ region. Further, the large size of the V region and the differences in recombination timing and dependence on the IL7R, Pax5, Ezh2 and YY1 in different domains, suggest that there may be boundary elements separating different regions. Furthermore there is 90kb of uncharacterized sequence between the last V_H and first D_H gene and it will be interesting to see if it contains any enhancers, or insulator elements to prevent the V region recombining before the DJ region.

Acknowledgements

The authors thank members of the Corcoran group for helpful discussions. This work was supported by the Biotechnology and Biological Sciences Research Council.

References

1. Hesslein DG, Schatz DG. Factors and forces controlling V(D)J recombination. Adv Immunol 2001; 78:169-232.
2. Corcoran AE. Immunoglobulin locus silencing and allelic exclusion. Semin Immunol 2005; 17(2):141-154.
3. Johnston CM, Wood AL, Bolland DJ et al. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. J Immunol 2006; 176(7):4221-4234.
4. Brekke KM, Garrard WT. Assembly and analysis of the mouse immunoglobulin kappa gene sequence. Immunogenetics 2004; 56(7):490-505.
5. Ye J. The immunoglobulin IGHD gene locus in C57BL/6 mice. Immunogenetics 2004; 56(6):399-404.
6. Atkinson MJ, Michnick DA, Paige CJ et al. Ig gene rearrangements on individual alleles of Abelson murine leukemia cell lines from (C57BL/6 × BALB/c) F1 fetal livers. J Immunol 1991; 146(8):2805-2812.
7. Jeong HD, Komisar JL, Kraig E et al. Strain-dependent expression of VH gene families. J Immunol 1988; 140(7):2436-2441.
8. Malynn BA, Yancopoulos GD, Barth JE et al. Biased expression of JH-proximal VH genes occurs in the newly generated repertoire of neonatal and adult mice. J Exp Med 1990; 171(3):843-859.
9. Matsuda F, Ishii K, Bourvagnet P et al. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. J Exp Med 1998; 188(11):2151-2162.
10. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 1998; 94(4):463-470.
11. Jones JM, Gellert M. The taming of a transposon: V(D)J recombination and the immune system. Immunol Rev 2004; 200:233-248.
12. Retter I, Chevillard C, Scharfe M et al. Sequence and Characterization of the Ig Heavy Chain Constant and Partial Variable Region of the Mouse Strain 129S1. J Immunol 2007; 179(4):2419-2427.
13. McBlane F, Boyes J. Stimulation of V(D)J recombination by histone acetylation. Curr Biol 2000; 10(8):483-486.
14. Mattick JS. The functional genomics of noncoding RNA. Science 2005; 309(5740):1527-1528.
15. Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. Nature 1985; 318(6045):475-478.
16. Thompson A, Timmers E, Schuurman RK et al. Immunoglobulin heavy chain germ-line JH-C mu transcription in human precursor B-lymphocytes initiates in a unique region upstream of DQ52. Eur J Immunol 1995; 25(1):257-261.
17. Reth MG, Alt FW. Novel immunoglobulin heavy chains are produced from DJH gene segment rearrangements in lymphoid cells. Nature 1984; 312(5993):418-423.
18. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. Cell 1985; 40(2):271-281.
19. Corcoran AE, Riddell A, Krooshoop D et al. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature 1998; 391(6670):904-907.
20. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
21. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85(6):887-897.
22. Bolland DJ, Wood AL, Johnston CM et al. Antisense intergenic transcription in V(D)J recombination. Nat Immunol 2004; 5(6):630-637.
23. Osborne CS, Chakalova L, Mitchell JA et al. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol 2007; 5(8):e192.
24. Fraser P. Transcriptional control thrown for a loop. Curr Opin Genet Dev 2006; 16(5):490-495.
25. Johnson K, Pflugh DL, Yu D et al. B-cell-specific loss of histone 3 lysine 9 methylation in the V(H) locus depends on Pax5. Nat Immunol 2004; 5(8):853-861.
26. Gribnau J, Diderich K, Pruzina S et al. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. Mol Cell 2000; 5(2):377-386.
27. Drewell RA, Bae E, Burr J et al. Transcription defines the embryonic domains of cis-regulatory activity at the Drosophila bithorax complex. Proc Natl Acad Sci USA 2002; 99(26):16853-16858.
28. Masternak K, Peyraud N, Krawczyk M et al. Chromatin remodeling and extragenic transcription at the MHC class II locus control region. Nat Immunol 2003; 4(2):132-137.
29. Bernstein BE, Kamal M, Lindblad-Toh K et al. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 2005; 120(2):169-181.
30. Wilson CJ, Chao DM, Imbalzano AN et al. RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodeling. Cell 1996; 84(2):235-244.
31. Cho H, Orphanides G, Sun X et al. A human RNA polymerase II complex containing factors that modify chromatin structure. Mol Cell Biol 1998; 18(9):5355-5363.

32. Wittschieben BO, Otero G, de Bizemont T et al. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. Mol Cell 1999; 4(1):123-128.
33. Krogan NJ, Kim M, Tong A et al. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol Cell Biol 2003; 23(12):4207-4218.
34. Ng HH, Robert F, Young RA et al. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol Cell 2003; 11(3):709-719.
35. Schwartz BE, Ahmad K. Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev 2005; 19(7):804-814.
36. Mito Y, Henikoff JG, Henikoff S. Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet 2005.
37. Orphanides G, Reinberg D. RNA polymerase II elongation through chromatin. Nature 2000; 407(6803):471-475.
38. Chakalova L, Debrand E, Mitchell JA et al. Replication and transcription: shaping the landscape of the genome. Nat Rev Genet 2005; 6(9):669-677.
39. Mahy NL, Perry PE, Bickmore WA. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. J Cell Biol 2002; 159(5):753-763.
40. Muller WG, Walker D, Hager GL et al. Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. J Cell Biol 2001; 154(1):33-48.
41. Volpi EV, Chevret E, Jones T et al. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J Cell Sci 2000; 113 (Pt 9):1565-1576.
42. Chambeyron S, Da Silva NR, Lawson KA et al. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development 2005; 132(9):2215-2223.
43. Schmitt S, Prestel M, Paro R. Intergenic transcription through a polycomb group response element counteracts silencing. Genes Dev 2005; 19(6):697-708.
44. Ho Y, Elefant F, Liebhaber SA et al. Locus control region transcription plays an active role in long-range gene activation. Mol Cell 2006; 23(3):365-375.
45. Sleutels F, Zwart R, Barlow DP. The noncoding Air RNA is required for silencing autosomal imprinted genes. Nature 2002; 415(6873):810-813.
46. Grewal SI, Jia S. Heterochromatin revisited. Nat Rev Genet 2007; 8(1):35-46.
47. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 2007; 8(4):272-285.
48. Katayama S, Tomaru Y, Kasukawa T et al. Antisense transcription in the mammalian transcriptome. Science 2005; 309(5740):1564-1566.
49. Rinn JL, Kertesz M, Wang JK et al. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. Cell 2007; 129(7):1311-1323.
50. Uhler JP, Hertel C, Svejstrup JQ. A role for noncoding transcription in activation of the yeast PHO5 gene. Proc Natl Acad Sci USA 2007; 104(19):8011-8016.
51. Sessa L, Breiling A, Lavorgna G et al. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. RNA 2007; 13(2):223-239.
52. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J 2001; 20(22):6394-6403.
53. Maes J, O'Neill LP, Cavelier P et al. Chromatin remodeling at the Ig loci prior to V(D)J recombination. J Immunol 2001; 167(2):866-874.
54. Johnson K, Angelin-Duclos C, Park S et al. Changes in histone acetylation are associated with differences in accessibility of V(H) gene segments to V-DJ recombination during B-cell ontogeny and development. Mol Cell Biol 2003; 23(7):2438-2450.
55. Su IH, Basavaraj A, Krutchinsky AN et al. Ezh2 controls B-cell development through histone H3 methylation and Igh rearrangement. Nat Immunol 2003; 4(2):124-131.
56. Bolland DJ, Wood AL, Afshar R et al. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. Mol Cell Biol 2007; 27(15):5523-5533.
57. Sakai E, Bottaro A, Davidson L et al. Recombination and transcription of the endogenous Ig heavy chain locus is effected by the Ig heavy chain intronic enhancer core region in the absence of the matrix attachment regions. Proc Natl Acad Sci USA 1999; 96(4):1526-1531.
58. Serwe M, Sablitzky F. V(D)J recombination in B-cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. EMBO J 1993; 12(6):2321-2327.
59. Perlot T, Alt FW, Bassing CH et al. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci USA 2005; 102(40):14362-14367.
60. Afshar R, Pierce S, Bolland DJ et al. Regulation of IgH Gene Assembly: Role of the Intronic Enhancer and 5'DQ52 Region in Targeting DHJH Recombination. J Immunol 2006; 176(4):2439-2447.

61. Morshead KB, Ciccone DN, Taverna SD et al. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci USA 2003; 100(20):11577-11582.
62. Maes J, Chappaz S, Cavelier P et al. Activation of V(D)J Recombination at the IgH Chain JH Locus Occurs within a 6-Kilobase Chromatin Domain and Is Associated with Nucleosomal Remodeling. J Immunol 2006; 176(9):5409-5417.
63. Bangs LA, Sanz IE, Teale JM. Comparison of D, JH and junctional diversity in the fetal, adult and aged B-cell repertoires. J Immunol 1991; 146(6):1996-2004.
64. Chakraborty T, Chowdhury D, Keyes A et al. Repeat organization and epigenetic regulation of the DH-Cmu domain of the immunoglobulin heavy-chain gene locus. Mol Cell 2007; 27(5):842-850.
65. Abarrategui I, Krangel MS. Regulation of T-cell receptor-alpha gene recombination by transcription. Nat Immunol 2006; 7(10):1109-1115.
66. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J 2001; 20(22):6394-6403.
67. Hesslein DG, Pflugh DL, Chowdhury D et al. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev 2003; 17(1):37-42.
68. Maes J, Chappaz S, Cavelier P et al. Activation of V(D)J recombination at the IgH chain JH locus occurs within a 6-kilobase chromatin domain and is associated with nucleosomal remodeling. J Immunol 2006; 176(9):5409-5417.
69. Kosak ST, Skok JA, Medina KL et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 2002; 296(5565):158-162.
70. Yang Q, Riblet R, Schildkraut CL. Sites that direct nuclear compartmentalization are near the 5' end of the mouse immunoglobulin heavy-chain locus. Mol Cell Biol 2005; 25(14):6021-6030.
71. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18(4):411-422.
72. Roldan E, Fuxa M, Chong W et al. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol 2005; 6(1):31-41.
73. Sayegh C, Jhunjhunwala S, Riblet R et al. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B-cells. Genes Dev 2005; 19(3):322-327.
74. Cobaleda C, Schebesta A, Delogu A et al. Pax5: the guardian of B-cell identity and function. Nat Immunol 2007; 8(5):463-470.
75. Liu H, Schmidt-Supprian M, Shi Y et al. Yin Yang 1 is a critical regulator of B-cell development. Genes Dev 2007; 21(10):1179-1189.
76. Iborra FJ, Pombo A, Jackson DA et al. Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. J Cell Sci 1996; 109 (Pt 6):1427-1436.
77. Faro-Trindade I, Cook PR. Transcription factories: structures conserved during differentiation and evolution. Biochem Soc Trans 2006; 34(Pt 6):1133-1137.
78. Osborne CS, Chakalova L, Brown KE et al. Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet 2004; 36(10):1065-1071.
79. Trimborn T, Gribnau J, Grosveld F et al. Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. Genes Dev 1999; 13(1):112-124.
80. Ragoczy T, Bender MA, Telling A et al. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev 2006; 20(11):1447-1457.
81. Chowdhury D, Sen R. Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements. Immunity 2003; 18(2):229-241.
82. Skok JA, Brown KE, Azuara V et al. Nonequivalent nuclear location of immunoglobulin alleles in B-lymphocytes. Nat Immunol 2001; 2(9):848-854.
83. Daly J, Licence S, Nanou A et al. Transcription of productive and nonproductive VDJ-recombined alleles after IgH allelic exclusion. EMBO J 2007; 26(19):4273-4282.
84. Haines BB, Brodeur PH. Accessibility changes across the mouse Igh-V locus during B-cell development. Eur J Immunol 1998; 28(12):4228-4235.
85. Skok JA, Gisler R, Novatchkova M et al. Reversible contraction by looping of the Tcra and Tcrb loci in rearranging thymocytes. Nat Immunol 2007; 8(4):378-387.
86. Mostoslavsky R, Singh N, Kirillov A et al. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. Genes Dev 1998; 12(12):1801-1811.
87. Goldmit M, Ji Y, Skok J et al. Epigenetic ontogeny of the Igk locus during B-cell development. Nat Immunol 2005; 6(2):198-203.
88. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128(4):693-705.
89. Pawlitzky I, Angeles CV, Siegel AM et al. Identification of a candidate regulatory element within the 5' flanking region of the mouse Igh locus defined by pro-B-cell-specific hypersensitivity associated with binding of PU.1, Pax5 and E2A. J Immunol 2006; 176(11):6839-6851.

CHAPTER 6

Genetic and Epigenetic Control of V Gene Rearrangement Frequency

Ann J. Feeney*

Abstract

The antibody repertoire is enormous and reflects the power of combinatorial and junctional diversity to generate a vast repertoire with a moderate number of V, D and J gene segments. However, although there are many V_H and $V\kappa$ gene segments, the usage of these genes is highly unequal. In this chapter, we summarize our studies elucidating many of the factors that contribute to this unequal rearrangement frequency of individual gene segments. Firstly, there is much natural variation in the sequence of the Recombination Signal Sequences (RSS) that flank each recombining gene. This genetic variation contributes greatly to unequal recombination frequencies. However, other factors also play a major role in recombination frequencies, as evidenced by the fact that some genes with identical RSS rearrange at very different frequencies in vivo. Analysis of these gene segments by chromatin immunoprecipitation (ChIP) suggests that differences in the structure of the chromatin associated with each gene is also a major factor in differential accessibility for rearrangement. Finally, transcription factors can direct accessibility for recombination, possibly by recruiting chromatin-modifying enzymes to the vicinity of the gene segment. Together, these factors dictate the composition of the newly formed antibody repertoire.

Introduction

The vast antibody repertoire is created by a combination of junctional diversity and combinatorial diversity. Each antibody heavy chain is encoded by a heavy chain and a light chain, the latter being encoded by either the kappa locus or the lambda locus. The heavy chain variable region is itself composed of three segments, V, D and J, while the light chain variable region has two segments, V and J. Combinatorial diversity is generated through the use of one each of the many V, D and J gene segments to encode the heavy and light chain exons and junctional diversity is generated by the deletion of a variable small number of nucleotides from the ends of each recombining gene segment and the random addition of a few nucleotides to the junction by TdT.[1] The BALB/c IgH locus contains ~50-100 functional V_H genes, 13 functional D_H genes and 4 functional J_H genes.[2-4] In the mouse, the random association of one V_H, one D_H and one J_H would theoretically create ~75 × 13 × 4 different H chains and the random association of light chain gene segments would similarly create ~50-100 $V\kappa$ × 4 $J\kappa$ kappa light chains and 4 different lambda chains. Further random association of heavy and light chains would thus create over a million different antibodies on the basis of combinatorial diversity alone. This theoretical diversity of combinations of gene segments has been thought to be a major factor in the size of the repertoire and given the large number of gene segments in the Ig loci, combinatorial diversity does contribute greatly to the antibody repertoire. However, we and others have shown that the rearrangement frequency

*Ann J. Feeney—The Scripps Research Institute, Department of Immunology, IMM22, 10550 N. Torrey Pines Rd, La Jolla, CA 92037, USA. Email: feeney@scripps.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

of the different gene segments is very unequal and thus the contribution of some V gene segments to the repertoire is much smaller that that of other V genes.[5-8] Some of the genetic and epigenetic reasons for this difference in gene usage are summarized here.

Sequence Variation in RSS Can Greatly Affect Recombination

Each gene segment is flanked by a recombination signal sequence (RSS) which is composed of a conserved heptamer and nonamer, separated by a spacer of ~12 or ~23 bp.[9,10] The heptamer and nonamer have consensus sequences, but there is great natural variation in the sequences found in the Ig and TCR loci. The pioneering work of Gellert and colleagues using plasmid-based recombination substrates containing two RSS which can be varied in sequence clearly showed that the sequence of the heptamer and nonamer of the RSS were very important in determining the frequency of recombination.[9] By varying a nucleotide at each position of the heptamer and nonamer and varying the spacer length, general rules were established which showed that the first three bp of the heptamer were critical and changes in those positions away from the consensus almost abolished recombination. In contrast, variation in other positions showed a wide range of decreased recombination. These studies were complemented by the RSS database analysis by Ramsden and Wu of all published Ig and TCR RSS as of 1994.[11] They demonstrated that the first three base pairs of the RSS, CAC, were essentially invariant, whereas other positions of the heptamer and nonamer had more variability.

In order to assess whether the natural variation in RSS could be responsible for the unequal rearrangement frequency, we first identified the frequency with which specific V genes rearranged in vivo before any biological selection could occur. We analyzed rearrangement of murine V_H genes in μMT mice, in which the mutation in the transmembrane exon of the heavy chain prevented differentiation past the pro-B-cell stage and we also analyzed rearrangement of Vκ genes in human cord blood cells.[7,8,12-15] In both cases we identified which genes rearranged more often than others in vivo. Then, using a modification of the recombination substrate approach, we determined if the RSS could be responsible for this nonrandom rearrangement. We designed "competition recombination substrates" in which, for example, two Vκ genes competed for rearrangement to a Jκ gene, as shown in Figure 1.[13] In this way, small differences in recombination could be assayed by determining the relative frequency with which the Jκ gene rearranged to each of the two Vκ genes. Each of the RSS in our plasmids were made by PCR so that they included ~50-100 bp of flanking DNA on either side of the RSS.

The analysis of two Vκ alleles provides a clear demonstration of the ability of a single base pair in the RSS to significantly affect recombination frequency. The VκA2 gene is used in the majority of anti-*Haemophilus influenzae* Type b (Hib) antibodies.[16] Navajos and genetically related Native

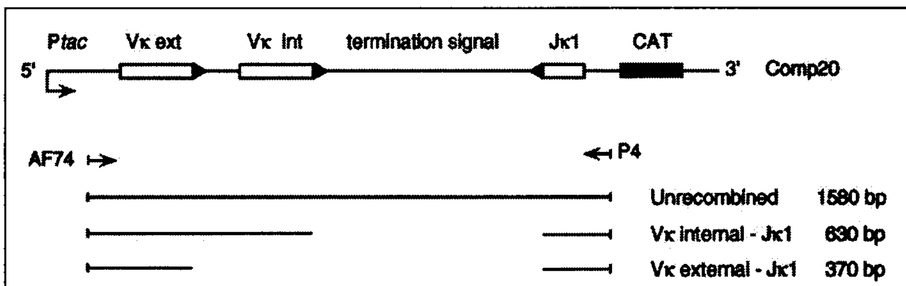


Figure 1. Competition recombination substrate. The top panel shows the basic design of the plasmid-based recombination substrate and the bottom panel shows the PCR assay used to determine the relative rearrangement of the Jκ gene to the internal or external Vκ. This basic design was used for all of our studies on the efficiency of various RSS. This figure is reproduced with permission from the Journal of Experimental Medicine, 1998, 187:1495-1503. Copyright 1998, The Rockefeller University Press.

Americans have a high incidence of Hib disease[17,18] and we discovered that they had a unique allele of the VκA2 gene, with one change from the predominant VκA2a allele at the 6th position in the heptamer.[15] Peripheral blood DNA from VκA2$^{a/b}$ heterozygotes showed that the VκA2a allele was rearranged ~5 times more often than this new Navajo VκA2b allele.[15] By placing the two VκA2 alleles in competition for a Jκ gene RSS in a recombination substrate, we were able to show that this single base pair change in the RSS was responsible for the difference in rearrangement frequency.[13] In this particular case, we hypothesized that this single nucleotide polymorphism was likely to play an important role in the increased incidence of Hib disease in Navajos, since impaired rearrangement of this Vκ gene would decrease the frequency of protective anti-Hib antibodies.[15] This would be one of the rare situations in which there was a genetic "hole" in the antibody repertoire with severe biological outcome: susceptibility to potentially fatal Hib disease.

RSS Is Not Always Responsible for Unequal Rearrangement

We found other examples where the rearrangement frequency in vivo was also recapitulated in the recombination substrate, demonstrating that the genetic basis for rearrangement differences was due to changes in the sequence of the RSS. For example, the small V_HS107 family has 3 functional V_H genes that rearrange at very different frequencies in vivo.[12] In pro-B-cells, the V1 gene rearranges 5 times more often than V11 and 40 times more often than V13. Using competition recombination substrates, we demonstrated that the V1 gene has an RSS that supports 3 times more rearrangement than the V11 RSS, thus accounting in large measure for the difference in rearrangement frequency in vivo.[14] However, V11 and V13 have very different rearrangement frequencies in vivo, yet their RSS are identical. Recombination substrate assays with ~100 bp fragments of V11 and V13 showed that the 5' and 3' DNA flanking the RSS also did not affect recombination frequency.[14] Hence, factors other than the RSS control the rearrangement frequency of these two V genes, as will be discussed later in this chapter.

In another example of genes with identical RSS rearranging at different frequencies, we analyzed the 20-member V_H7183 gene family. This is the most proximal V_H family, along with the V_HQ52 family that is interspersed with it in the 250 kb at the 3' end of the V_H locus. The most 3' functional V_H gene in this family, 81X, has been shown by several groups to rearrange at an extremely high frequency,[19,20] but the frequency of rearrangement of the other members of the family had not been determined. We analyzed the rearrangement frequency of the entire V_H7183 family in pro-B-cells and showed that the genes rearranged with a wide range of frequencies.[8] We cloned and sequenced each of the genes in the family and the RSS fell into two major groups. One group, which we termed Group I, had an RSS that was closer to the consensus than Group II and in competition recombination substrates, we showed that the Group I RSS supported a higher frequency of rearrangement than the Group II RSS, as would be predicted.[8] However, the rearrangement frequency of V_H genes with identical RSS was quite different in vivo in many cases and the Group I genes did not rearrange at a higher frequency than the Group II genes. Thus, factors other than the RSS were more important than the differences in the efficiency of the RSS in controlling gene rearrangement frequency for this V_H gene family. We mapped all of the V_H genes in the family and found a much higher correlation between chromosomal location and V gene rearrangement frequency.[8] The genes closest to 81X at the 3' portion of the locus rearranged more than the V_H genes in the middle of the locus and the genes in the 5' third of the locus rearranged very poorly, with the exception of the last V_H gene in the family, 61-1P. We propose that the chromatin structure may be different at these different portions of the V_H7183 part of the V_H locus, resulting in the observed different rearrangement frequencies for genes with identical RSS scattered throughout this 250 kb region.

Chromatin as the Gatekeeper of Accessibility

The process of gene rearrangement is lineage-specific, in that TCR genes do not rearrange in B-cells and Ig genes do not rearrange in T-cells, other than some D_H-J_H rearrangements.[21] Furthermore, this process of V(D)J rearrangement is highly ordered: D_H to J_H, followed by V_H to

DJ_H, followed by kappa rearrangement and lastly lambda rearrangement. The same order is observed in T-cells, with TCRβ rearrangement occurring before TCRα. Over 2 decades ago, Alt and colleagues proposed the "accessibility hypothesis" to explain these observations.[22] This hypothesis stated that accessibility to recombination would be limited to only certain small portions of the Ig or TCR loci in any given lymphocyte precursor population, e.g., the D_H and J_H subloci in early pro-B-cells. This hypothesis was supported by the observation that germline transcription of unrearranged genes precedes gene rearrangement, thus suggesting that this transcription reflected the induced accessibility for RAG binding and rearrangement.[23] The mechanism by which regions were maintained in inaccessible status until the proper time for their rearrangement was not clear at that time, but it is now generally agreed that chromatin structure is likely to be the key factor.[24]

The tails of histone proteins protrude from the core nucleosome and they can be posttranslationally modified by acetylation, methylation, phosphorylation and ubiquitinylation.[25,26] In general, lysines on tails of histones H3 and H4 are acetylated on active genes. Methylations are more complex and methylation of specific lysines, such as lysine 9 (K9me) or lysine 27 (K27me) on H3, are associated with repressed genes in general, while methylation of lysine 4 (K4me) on H3 is associated with active genes. It has been shown that V, D and J genes that are rearranging are more highly associated with acetylated H3 and H4 and less frequently associated with repressive modifications such as H3K9me, than genes that are not rearranging at that particular stage in lymphocyte development.[27-30] Thus, the status of histone posttranslational modifications (PTM) may control the accessibility of V, D and J genes.

Since histone acetylation appears to affect the accessibility of V, D and J genes, we hypothesized that perhaps the V genes that did not rearrange as well were associated with histones that did not have as high an extent of this positive PTM and were higher in negative PTM such as H3K9me. We therefore analyzed the V_HS107 gene family by chromatin immunoprecipitation (ChIP) with antibodies against acetylated H3 and H4. We were particularly interested in determining if there were any differences between V11 and V13, since they had identical RSS yet rearranged at such different frequencies. Indeed, there was an excellent correlation between the relative rearrangement frequency of the three V_HS107 genes and their enrichment in acetylated H3 and H4 (Fig. 2).[31] Furthermore, there was an inverse relationship between the level of the repressive modification

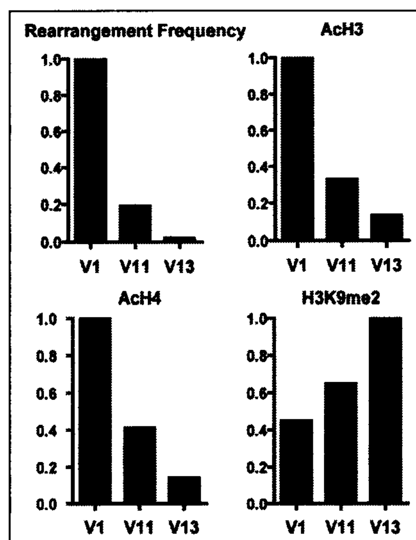


Figure 2. Relative rearrangement frequency in vivo of the three functional V_HS107 genes correlates positively with the extent of histone acetylation and negatively with the extent of histone K9 methylation.

H3K9me2 and the rearrangement frequency (Fig. 2).[32] Thus, the histone PTM status accurately reflects the relative accessibility for rearrangement of these genes.

We further investigated the histone PTM status of the 81X gene. This gene rearranges at a very high frequency in fetal life and also rearranges at a high frequency in adult bone marrow, although not as frequently as in fetal liver. We compared the histone acetylation status of this gene as compared to the rest of the V_H7183 family, using an 81X-specific primer and a primer that amplifies all V_H7183 genes except 81X. 81X was more highly enriched in acetylated histones than the remainder of the V_H7183 family and the extent of enrichment was greater in fetal life than in adult life, correlating with the relatively higher rearrangement of 81X in fetal life (Fig. 3).[31] As with the V_HS107 genes, there was a reciprocal relationship of these genes with the repressive H3K9me2 PTM.[32]

Further evidence that histone PTM may influence rearrangement frequency is demonstrated by analysis of mice deficient in the histone methyltransferase Ezh2, which adds the repressive K27 methylation PTM. The pro-B-cells from these mice do not rearrange the V_H genes in the distal half of the locus, although the proximal half rearranges at near normal frequency.[33] We have shown that the H3K27me PTM is found on the proximal V_H genes in pro-B-cells (C.-R. Xu and AJF, unpublished data) and thus we propose that the presence of this repressive PTM on the proximal V_H genes is necessary for the distal V_H genes to rearrange at normal frequency.

Role of Transcription Factors in Controlling Rearrangement

Although there clearly seems to be a good correlation between histone PTM patterns and accessibility for recombination, it is not clear what determines the histone modification status of genes. Histone acetylases, deacetylases and methylases are often recruited into large multi-protein complexes and it is likely that the specificity of these complexes derives from DNA-binding transcription factors. We have investigated the role of transcription factors in inducing accessibility for recombination. Mice that are deficient in EBF, Pax5, or E2A, are all devoid in B-cells, demonstrating the essential role of these transcription factors in B-cell differentiation.[34-37] The first two factors are B-cell specific and are essential for B-cell development. E2A is a widely expressed transcription factor, but it is only in B-cells that it is present as a homodimer and this probably explains the specific loss of B-cells in the E2A-deficient mice.[38]

Using a novel system, devised by our collaborator Dr Cornelis Murre in which expression vectors for E2A or EBF were transiently transfected, along with expression vectors for RAG1 and RAG2, into a nonlymphoid cell line, the ability of transcription factors to induce accessibility of genes for recombination was revealed.[39] Transient transfection with EBF resulted in the induction of rearrangement of Vλ3 genes, but not of any kappa genes. Conversely, the ectopic expression of E2A resulted in recombination of many VκI genes in this cell line. Importantly, although the three major Vκ

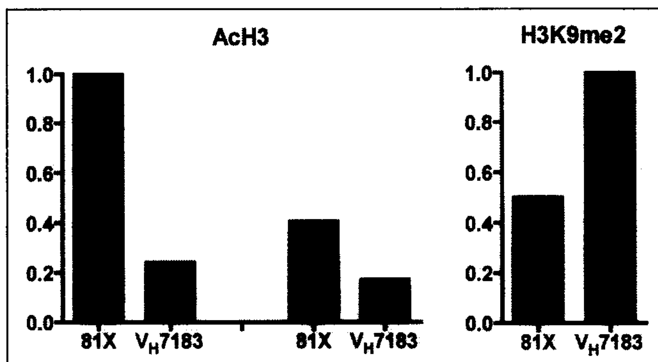


Figure 3. The frequently rearranging 81X gene is more highly enriched for histone acetylation and less enriched for H3K9 methylation than the remainder of the V_H7183 family.

families are interspersed, only the VκI genes are induced to rearrange, but not the neighboring VκII or VκIII genes (Fig. 4).[40] Thus, this means that the Vκ locus is not made accessible as a whole unit by the action of E2A, but that individual V genes, or related V genes such as members of a Vκ family, are induced on a localized level to become accessible. Since members of a Vκ or V_H family arose by gene duplication, their coding and flanking sequences are very similar. Therefore, we proposed that there are transcription factor binding sites in the vicinity of all functional V genes and that the binding of the appropriate transcription factor could then recruit chromatin modifying enzymes such as histone acetyltransferases or deacetylases, histone methyltransferases or demethylases, or ATP-dependent chromatin remodeling complexes, which would then change the chromatin structure of the V gene, making it accessible, or inaccessible, for recombination.[40]

We hypothesized that the expression of E2A would increase the histone acetylation of the VκI genes, but not of the VκII and VκIII genes which were not induced to undergo recombination after ectopic E2A expression. Similarly, we hypothesized that expression of EBF would increase the histone acetylation of Vλ3 genes specifically. We assessed this by ChIP, using primers that flanked the RSS and we found that this was indeed the case (P Goebel and AJF, unpublished data). Surprisingly, however, we found that the extent of acetylation of the appropriate genes was very modest. We

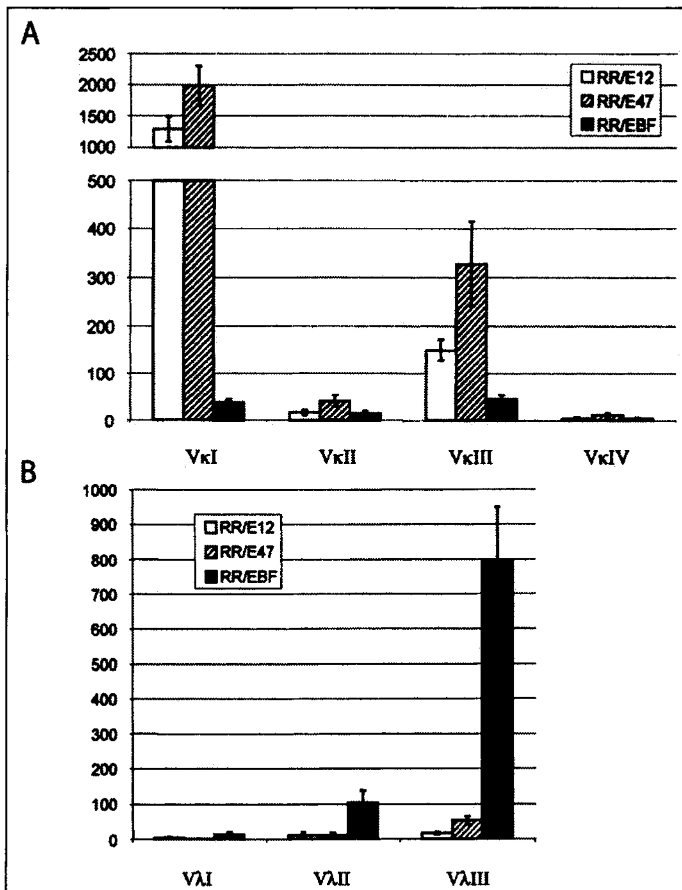


Figure 4. Ectopic expression of E2A in a nonlymphoid cell line induces preferential rearrangement of VκI genes (A), while ectopic expression of EBF induces preferential rearrangement of Vλ3I (B). This figure is reproduced with permission from the *Journal of Experimental Medicine*, 2001, 194:645-656. Copyright 2001, The Rockefeller University Press.

therefore proposed that only a small fraction of the Vκ genes are induced to become acetylated by E2A and that these genes are preferentially used for recombination. Similarly, our data showing that within the V_HS107 family, the V1 gene is most highly enriched in acetylated histones and V13 the least, most likely reflects the fact that more V1 genes within the population of pro-B-cells that we investigated were associated with acetylated histones.[31] This may suggest that the limiting factor for recombination is the number of V genes that are acetylated at any given time.

Pax5 also has an important role in V(D)J recombination, in addition to its role in controlling the expression of hundreds of genes critical for B-cell function.[41] Mice deficient in Pax5 are blocked at the late pro-B-cell stage of differentiation.[42] Although the proximal V_H7183 family rearranges at almost normal frequency, the distal V_HJ558 genes seldom rearrange and the V_H families in between these two families rearrange at intermediate levels.[43] An explanation for this could be that the V_H locus in Pax5-deficient pro-B-cells fails to undergo the compaction that appears to be critical to bring the distal V_H genes closer to the D-J_H locus.[44] In this extended configuration, the distal V_HJ558 genes, which are located 1-2.5 Mb from the D-J region, would be too far away from the DJ_H genes to undergo rearrangement. In addition, it has been shown that V_H genes in Pax5-deficient B-cells are enriched in the repressive modification H3K9me2 and it has been suggested that Pax5 is required for the histone exchange necessary to make the V_H genes associate with acetylated histone H3 and not K9 methylated H3.[45]

We have described another function for Pax5 which is important for V(D)J recombination. Although transcription factor binding sites are traditionally found in promoters and enhancers, we searched for Pax5 binding sites within V_H coding regions since we hypothesized that the RAG complex may bind Pax5. The reason for this hypothesis was the fact that the core RAG2 knock-in mice had a defect in V to DJ rearrangement, although DJ and kappa rearrangement was not impaired.[46] Since Pax5-deficient mice were originally reported to have the same generalized defect in V_H rearrangement, but not DJ rearrangement,[42] we hypothesized that perhaps the non-core region of RAG2 might bind to Pax5 which would stabilize its interaction with the RSS. A search of V_H gene sequences with the sequences of the Pax5 binding sites in KI, KII, RAG, CD19, showed several potential matches (AFJ, unpublished data) and EMSA analysis showed Pax5 did bind to several of these sites, with varying affinities.[32,47] The V_HS107 gene V1 had the highest affinity site and the V_H7183 genes had strong Pax5 binding sites also. The V_HJ558 genes also had Pax5 binding sites, although their affinity estimated by cold target competition was lower. ChIP analysis showed that Pax5 was bound to V_H genes in pro-B-cells.[47] Our collaborators, Zhixin Zhang and Max Cooper, showed that Pax5 interacted with RAG complex, although they showed that Pax5 also bound to complexes made with the core RAG1/2,[47] thus rendering our initial hypothesis that Pax5 may bind to the non-core portion of RAG2 unlikely. Using an in vitro assay, our collaborators found that Pax5 increased recombination, suggesting that the interaction of Pax5 with the RAG complex did stabilize the interaction.[47] In addition to this role of Pax5 in interacting with the RAG complex, we propose that these Pax5 sites located throughout the *IgHV* locus may be the reason that Pax5-deficient mice cannot undergo locus contraction and thus the function of these Pax5 sites in V_H genes may be to initiate *IgHV* locus contraction. Mice deficient in the transcription factor YY1 also have a defect in rearranging distal V_H genes and also do not undergo locus compaction[48] and thus a complex containing Pax5 and YY1 complex may be involved in the contraction of the locus.

Conclusion

The antibody repertoire derives part of its size from the combinatorial diversity generated when different V, D and J genes are used to encode the two chains of the receptor heterodimers. However, all V, D and J genes are used at very different frequencies. We have summarized work showing that part of this unequal representation is due to the natural variation in the sequences of the RSS flanking each gene. Since the RSS is the DNA binding site for the RAG recombinase, the mechanism for the influence of these genetic variations is clear. However, the differences in the chromatin structure of nucleosomes associated with individual V genes can override the simple

direct effect of the genetic variation in RSS efficiency in recruiting and stabilizing RAG binding. We found that the extent of positive or negative histone PTM can affect the ability of individual V genes to undergo rearrangement. One of the important unanswered questions is to determine what directs the histone modifications to occur on specific genes within the V loci. We hypothesize that specific transcription factors bind to sites near the V genes, in the promoter or even in coding regions as we have shown for Pax5. These proteins may then recruit histone-modifying enzymes, chromatin remodeling complexes and DNA methyltransferases. These epigenetic modifications would then render a gene more or less accessible or inaccessible to undergo rearrangement.

Acknowledgements

This was supported by NIH grants R01 AI29672 and AI52313. I gratefully acknowledge the many contributions of present and former lab members to this research and in particular would like to acknowledge the work of Drs Cheng-Ran Xu, Bertrand Nadel, Peter Goebel, Celia Espinoza and G Stuart Williams.

References

1. Tonegawa S. Somatic generation of antibody diversity. Nature 1983; 14:575-581.
2. Retter I, Chevillard C, Scharfe M et al. Sequence and characterization of the Ig heavy chain constant and partial variable region of the mouse strain 129S1. J Immunol 2007; 179(4):2419-2427.
3. Johnston CM, Wood AL, Bolland DJ et al. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. J Immunol 2006; 176(7):4221-4234.
4. Feeney AJ, Riblet R. D_{ST4}: a new and probably the last, functional D_H gene in the BALB/c mouse. Immunogenetics. 1993; 37:217-221.
5. Brezinschek HP, Brezinschek RI, Lipsky PE. Analysis of the heavy chain repertoire of human peripheral B-cells using single-cell polymerase chain reaction. J Immunol 1995; 155:190-202.
6. Foster SF, Brezinschek H-P, Brezinschek RI et al. Molecular mechanisms and selective influences that shape the kappa gene repertoire of IgM^+ cells. J Clin Invest 1997; 99:1614-1627.
7. Feeney AJ, Lugo G, Escuro G. Human cord blood κ repertoire. J Immunol 1997; 158:3761-3768.
8. Williams GW, Martinez A, Montalbano A et al. Unequal V_H gene rearrangement frequency within the large V_H7183 gene family is not due to RSS variation and mapping of the genes shows a bias of rearrangement based on chromosomal location. J Immunol 2001; 167:257-263.
9. Hesse JE, Lieber MR, Mizuuchi K et al. V(D)J recombination: A functional definition of the joining signals. Genes Dev 1989; 3:1053-1061.
10. Akira S, Okazaki K, Sakano H. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. Science 1987; 238:1134-1138.
11. Ramsden DA, Baetz K, Wu GE. Conservation of sequence in recombination signal sequence spacers. Nucl Acids Res 1994; 22:1785-1796.
12. Love VA, Lugo G, Merz D et al. Individual promoters vary in strength, but the frequency of rearrangement of those V_H genes does not correlate with promoter strength nor enhancer independence. Molecular Immunol 2000; 37:29-39.
13. Nadel B, Tang A, Lugo G et al. Decreased frequency of rearrangement due to the synergistic effect of nucleotide changes in the heptamer and nonamer of the recombination signal sequence of the V kappa gene A2b, which is associated with increased susceptibility of Navajos to Haemophilus influenzae Type b disease. J Immunol 1998; 161(11):6068-6073.
14. Montalbano A, Ogwaro KM, Tang A et al. V(D)J recombination frequencies can be profoundly affected by changes in the spacer sequence. J Immunol 2003; 171(10):5296-5304.
15. Feeney AJ, Atkinson MJ, Cowan MJ et al. A defective VκA2 allele in Navajos which may play a role in increased susceptibility to Haemophilus influenzae Type b disease. J Clin Invest 1996; 97:2277-2282.
16. Scott MG, Crimmins DL, McCourt DW et al. Clonal characterization of the human IgG antibody repertoire to Haemophilus influenzae Type b polysaccharide III. A single VκII gene and one of several Jκ genes are joined by an invariant arginine to form the most common L chain V region. J Immunol 1989; 143:4110-4116.
17. Coulehan JL, Michaels RH, Hallowell C et al. Epidemiology of Haemophilus influenzae Type b disease among Navajo indians. Pub Hlth Rep 1984; 99:404.
18. Petersen GM, Silimperi DR, Rotter JI et al. Genetic factors in Haemophilus influenzae Type b disease susceptibility and antibody acquisition. J Pediatrics 1987; 110:228-233.
19. Perlmutter RM, Kearney JF, Chang SP et al. Developmentally controlled expression of immunoglobulin V_H genes. Science 1985; 227:1597-1600.

20. Yancopoulos GD, Desiderio SV, Paskind M et al. Preferential utilization of the most J_H-proximal V_H gene segments in pre-B-cell lines. Nature 1984; 311:727-733.
21. Alt FW, Blackwell TK, Yancopoulos GD. Development of the primary antibody repertoire. Science 1987; 238:1079-1087.
22. Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. Ann Rev Immunol 1986; 4:339-368.
23. Yancopoulos GD, Malynn BA, Alt FW. Developmentally regulated and strain-specific expression of murine V_H gene families. J Exp Med 1988; 168:417-435.
24. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
25. Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape. Cell 2007; 128(4):635-638.
26. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128(4):693-705.
27. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J 2001; 20(22):6394-6403.
28. McMurry MT, Krangel MS. A role for histone acetylation in the developmental regulation of VDJ recombination. Science 2000; 287(5452):495-498.
29. Johnson K, Angelin-Duclos C, Park S et al. Changes in histone acetylation are associated with differences in accessibility of V_H gene segments to V-DJ recombination during B-cell ontogeny and development. Mol Cell Biol 2003; 23(7):2438-2450.
30. Morshead KB, Ciccone DN, Taverna SD et al. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci USA 2003; 100(20):11577-11582.
31. Espinoza CR, Feeney AJ. The extent of histone acetylation correlates with the differential rearrangement frequency of individual V_H genes in pro-B-cells. J Immunol 2005; 175:6668-6675.
32. Espinoza CR, Feeney AJ. Chromatin accessibility and epigenetic modifications differ between frequently and infrequently rearranging V(H) genes. Mol Immunol 2007; 44(10):2675-2685.
33. Su IH, Basavaraj A, Krutchinsky AN et al. Ezh2 controls B-cell development through histone H3 methylation and Igh rearrangement. Nat Immunol 2003; 4(2):124-131.
34. Urbanek P, Wang ZQ, Fetka I et al. Complete block of early B-cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. Cell 1994; 79(5):901-912.
35. Bain G, Maandag ECR, Izon D et al. E2A proteins are required for proper B-cell development and initiation of imunoglobulin gene rearrangements. Cell 1994; 79:885-892.
36. Lin H, Grosschedl R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature 1995; 376(6537):263-267.
37. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B-cell formation. Cell 1994; 79:875-884.
38. Bain G, Gruenwald S, Murre C. E2A and E2-2 are subunits of B-cell-specific E2-box DNA-binding proteins. Mol Cell Biol 1993; 13(6):3522-3529.
39. Romanow WJ, Langerak AW, Goebel P et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Mol Cell 2000; 5(2):343-353.
40. Goebel P, Janney N, Valenzuela JR et al. Localized gene-specific induction of accessibility to V(D)J recombination induced by E2A and early B-cell factor in nonlymphoid cells. J Exp Med 2001; 194(5):645-656.
41. Schebesta A, McManus S, Salvagiotto G et al. Transcription factor Pax5 activates the chromatin of key genes involved in B-cell signaling, adhesion, migration and immune function. Immunity 2007; 27(1):49-63.
42. Nutt LS, Urbanek P, Rolink A et al. Essential functions of Pax5 (BSAP) in pro-B-cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. Genes Dev 1997; 11:476-491.
43. Hesslein DG, Pflugh DL, Chowdhury D et al. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev 2003; 17(1):37-42.
44. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18(4):411-422.
45. Johnson K, Pflugh DL, Yu D et al. B-cell-specific loss of histone 3 lysine 9 methylation in the V_H locus depends on Pax5. Nat Immunol 2004; 5(8):853-861.
46. Liang HE, Hsu LY, Cado D et al. The "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B and T-cell development. Immunity 2002; 17(5):639-651.
47. Zhang Z, Espinoza CR, Yu Z et al. Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V_H to-DJ_H rearrangement of immunoglobulin genes. Nat Immunol 2006; 7(6):616-624.
48. Liu H, Schmidt-Supprian M, Shi Y et al. Yin Yang 1 is a critical regulator of B-cell development. Genes Dev 2007; 21(10):1179-1189.

Chapter 7

Dynamic Aspects of TCRα Gene Recombination:

Qualitative and Quantitative Assessments of the TCRα Chain Repertoire in Man and Mouse

Evelyne Jouvin-Marche,* Patrizia Fuschiotti and Patrice Noël Marche

Abstract

Most T-lymphocytes express a highly specific antigen receptor (TCR) on their cell surface, consisting of a clonotypic αβ-heterodimer. Both α- and β chains are products of somatic rearrangements of V, (D) and J gene segments encoded on the respective loci. The qualitative, quantitative and dynamic aspects of the TCRα chain repertoire of humans and mice have been difficult to estimate, mainly due to locus complexity. Analyses of the T-cell repertoire were first performed at the transcriptional level using classical cloning and sequencing strategies and then later at the genomic level using sensitive multiplex PCR assays that allow surveying the global rearrangement of the TCRAD locus. These all converge and support the conclusion that the V-J recombination pattern in both human and mouse thymus is not random but depends on the reciprocal V and J positions within the locus, thereby limiting the combinatorial diversity of the TCRα chain repertoire. The recombination profile is compatible with a sequential opening of the V region with progressive tracking along the two regions in opposite directions starting from the nearest and then moving towards the most distant V and J gene segments. In this chapter, we report new insights into the degree of human and mouse TCRα chain diversity in thymic and peripheral T-lymphocytes. Since the comparison of human and mouse V-J recombination shows a similar pattern of rearrangement, we suggest that spatial and temporal synchronization on the accessibility of V and J gene segments are general features of V-J rearrangements that are conserved throughout evolution.

Introduction

T-cell function relies on the specific recognition of foreign antigens. The majority of T-lymphocytes from humans and rodents express a clonotypic αβ TCR, which is a membrane-bound heterodimer composed of α and β chains that specifically respond to peptides derived from pathogens and bound to self-MHC molecules.[1] Each chain contains a constant domain and a variable domain, the latter being responsible for MHC and peptide recognition via interaction with highly diverse complementary-determining region (CDR) loops.[2] These chains are produced in differentiating lymphocytes by a series of somatic, site-specific DNA recombination reactions of multiple gene segments encoding TCR V, D and J domains.[3]

*Corresponding Author: Evelyne Jouvin-Marche—Université Joseph Fourier-Grenoble I, Faculté de Médecine, Institut d'Oncologie/Developpement Albert Bonniot et Institut Français du Sang, UMR-S823, Grenoble, France. Email: evelyne.marche@cnrs-dir.f

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

Lymphocytes have evolved sophisticated mechanisms for generating a diverse TCR repertoire. Multiple different copies of the V, (D) and J gene segments are each capable of contributing to a TCR antigen recognition domain and different combinations of gene segments can be used in independent rearrangement events. In addition to combinatorial diversity, variability is introduced by random removal and addition of nucleotides at the V-J or V-DJ junctions.[4] This nontemplated mechanism considerably increases the repertoire. A further diversifying factor is the pairing of α and β chains[5,6] to form TCR heterodimers. The potential diversity generated by random V(D)J recombination has been estimated at 10^{15} αβ TCRs.[2] However, this number is much higher than the actual size of the peripheral T-cell compartment, estimated at around 10^8 in mouse and 10^{12} in human. Furthermore, at least some cells express the same TCR specificity.[7,8] Consequently, at any given time, only a fraction of the potential repertoire (i.e., according to the random model) is achieved implying that other mechanisms must govern immune diversity.

In retrospect, the theoretical diversity of αβT-lymphocytes has been overestimated in several ways. Firstly, the T-cell repertoire has been evaluated assuming that any V gene can rearrange with any J gene in the TCRA locus. However, several sets of data on the mouse (thymus) model indicate that the number of V-J combinations is considerably lower due to a preferential association between V and J gene segments which depends on their position within the locus.[9-14] Secondly, the pairing of α and β chains to form the TCR heterodimer is constrained by structural compatibility between the subunits, further limiting the repertoire.[5,6] Thirdly, within the thymus, the newly generated repertoire is positively selected[15,16] via interactions with self MHC molecules expressed on stromal cells, reducing the size of the generated repertoire by approximately 100-fold. Furthermore, the establishment of a peripheral T-cell repertoire depends not only on the interactions of each T-cell with their respective ligands but also on complex homeostatic mechanisms ensuring the maintenance of numbers and immune functions of lymphocyte populations.[17]

Clearly, the size of the available peripheral TCRαβ diversity is difficult to determine and is open to debate. While the total number of lymphocytes in the blood can be measured directly, the diversity of the lymphocyte compartment on which immunocompetence is based cannot. Despite considerable knowledge of the determinants and profile of the TCRβ chain repertoire, very little is known about human and mouse TCRα chain diversity likely due to the TCRAD locus complexity and the limited number of anti-VAD antibodies available. Thus, we have only a partial view of the entire TCRA repertoire. Molecular measurements of TCR diversity using CDR3 length analysis[18] estimated about 0.5×10^6 different α chains and 10^6 different β chains expressed in human blood lymphocytes.[19] However, this calculation was based on the analysis of TCRβ transcripts expressed in αβ T-cell clones using some V genes and with the following two assumptions: 1) the probability of rearrangement between any V gene and J gene is equal; and 2) the V families are expressed at the same level.

Evaluation of the TCR repertoire is an important measure of the immunological competence of an individual. Animal models have been more extensively studied but the degree to which these results apply to the human model has yet to be established. By making comparisons between species, we hope to learn about the general principles in operation as well as their specific origins and what this may imply about the evolution of immunity.

Complexity of Mouse and Human TCRAD Locus

The maps of both mouse and human TCRAD loci have been elucidated in the last decade and are updated by IMGT.[20-22] Briefly, the human TCRAD locus spans about 1000 kb and consists of 54 V genes belonging to 41 families including 8 to 10 pseudogenes, 61 J gene segments, as well as 12 J pseudogenes, giving 49 functional Js and a unique C gene.[20,23-25] Similarly, the mouse TCRAD locus is composed of 70 to more than 100 V genes depending on the haplotype, regrouped into 23 families, 60 J gene segments including 16 pseudogenes (namely J1, 3, 4, 8, 14, 19, 20, 25, 29, 36, 41, 46, 51, 55, 59 and 60) giving 44 functional Js[14,20] and a unique C gene. In conclusion, the human J region contains more functional J segments able to rearrange than its mouse homologue (49 functional Js in human against 44 in mouse), providing more combinatory possibilities for the human V genes and compensating in part for the lower number of V genes compared to that in mice.

Analysis of Human and Mouse TCRA-Chain Diversity

Our previous studies on the V2 gene family of the mouse TCRAD locus indicated that rather than being stochastic, V2-J gene rearrangements depend on the respective location of the gene and occur in concentric waves.[12,26] During T-cell development, J usage moves from J genes which are the closest to the V gene region to J genes located farthest from this region; similarly, V2 usage moves from V2 genes closest to the J gene region to V2 genes located at the extremity of the locus. In other words, the most proximal V2 genes target the most proximal J gene segments whereas the most distal V2 genes rearrange preferentially with the most distal Js. However, these studies were focused on V2 genes and considered them representative of all V genes. Furthermore, the analysis of V2-J gene combinations was conducted at the mRNA level. One cannot therefore exclude varied transcription efficiency between different V2 genes that may affect the distribution of the V2-J combinations. To obtain a more accurate view of the V-J diversity, we must analyze all V-J combination events at the genomic level. As already mentioned, the diversity of the mammalian TCR repertoire is generated by gene rearrangement. We therefore developed a PCR assay allowing visualization at the DNA level of several contiguous recombination events between a given V gene or V gene family and several J genes segments of the TCRAD locus. As described in Figure 1, in each PCR assay, J primers were chosen to hybridize a downstream sequence allowing amplification of four to seven different J genes. Thus, a panel of nine to eleven J primers allowed the description of the rearrangement status of all functional mouse and human J genes and provided a global visualization of rearrangement patterns (Fig. 2).

Genomic multiplex PCR analysis of mouse TCRα chain diversity confirms previous data at the rearrangement level, in that V-J rearrangements are not random but depend on the V and J positions within the locus. For example, in the mouse thymus, V families located closest to the C coding region, such as V19 and V20, rearrange predominantly with the most proximal Js (J60 to J48) and rarely with the J segments located in the mid-section or the distal part of the J region (shown in Fig. 3). Reciprocally, V1 and V2 situated in the most distal part of the V gene region preferentially rearrange with the J segments found in the mid-section or distal parts of the J region but not with the Js found more proximally. Thus, the TCRAD locus is accessible from the 3′ end of the V region and from the 5′ end of the J region and consequently the proximal V and J genes are the first gene segments accessible for recombination followed later on by more distal V and J segments. In addition, we reported that depending on its locus position, each V gene differentially rearranged with a set of contiguous Js with a gaussian-like distribution.[14] For instance, the real time PCR quantification of V1 and V21 rearrangements revealed that the proximal V21 gene used a small set of J genes, less than 10, but with a 6 fold higher frequency than distal V genes which used a larger panel of J genes (more than 32). These preferential associations between V and J genes were observed with different V genes located at different positions in the TCRAD locus, suggesting that each V gene targeted particular sets of J segments.

A similar multiplex PCR experimental approach has been used to characterize the α chain repertoire in human thymi. By focusing the analysis on single member families to correlate the position of each V gene with its rearrangement pattern (Fig. 3, top panels), it can be observed that the two V genes most distant from the J region (V1, V2, located at -925 and -835 kb from the C gene, respectively) rearrange with the central and 3′ end of the J region, whereas the three J-proximal V genes, namely V38, V40 and V41, located between -267 to -227 kb with respect to the C gene, mainly rearrange with the most proximal Js. Finally, the members of the multigenic V8 family located in the middle part of the locus, including members located at -701, -653 and -569 kb respectively from the C region, rearrange to approximately the same extent with all the J segments throughout the locus. Taken together, the data show preferential distribution of recombination of particular V families to certain J gene segments depending on their localization within the locus. These findings are consistent with the model of synchronized waves of accessibility moving in a concentric manner across both V and J gene regions. These waves of rearrangement move from J genes located proximal to the V region towards J genes located closer to the C gene and from V genes located proximal to J region towards more distally located V genes, supporting the bi-directional

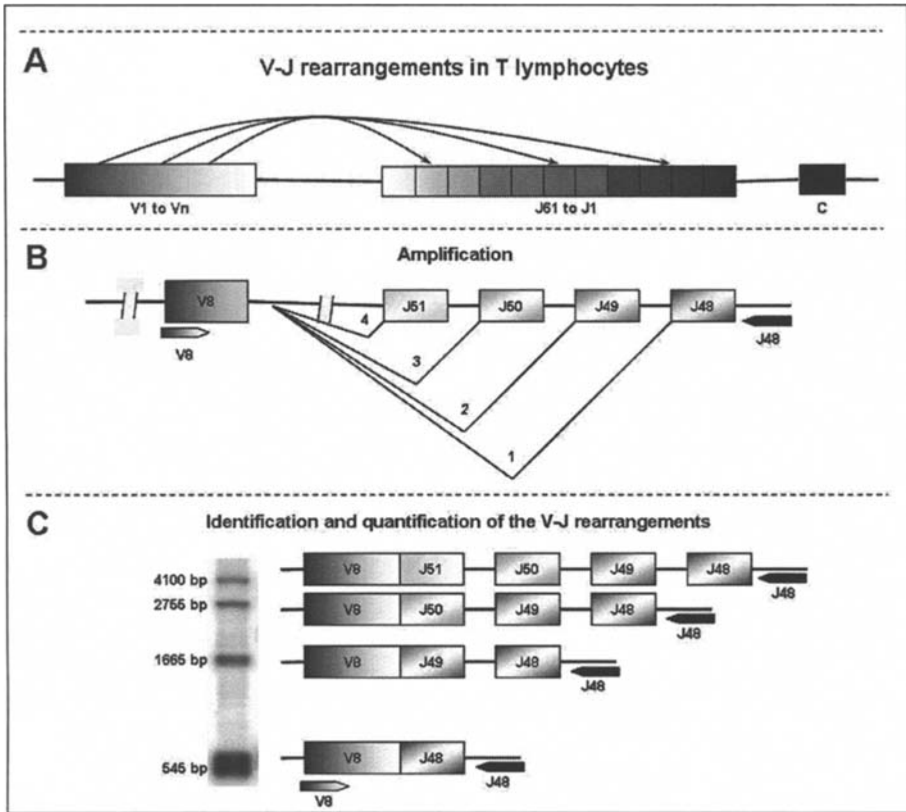


Figure 1. Schematic representation of multiplex PCR analysis of TCRA gene rearrangements. Briefly, by using two specific primers, one upstream of a given V gene and another downstream of a given J gene, the PCR will amplify all rearrangements involving both of these genes. This multiplex assay allows the detection of a V-J rearrangement as well as that of a set of four to seven upstream J genes with a maximum amplicon size of approximately 8 kb. The specificity of TCR rearrangement products can be assessed both by successive hybridization with internal V and J probes and by an accurate measure of the length of the PCR products compared to the known position of the genes in the locus.[21]

and coordinated model postulated in the mouse.[13,14] In conclusion, the comparison of human and mouse TCRA V-J recombination in the thymus shows a similar pattern of rearrangement suggesting that this mechanistic regulation of the process is conserved throughout evolution.

Comparison between the Frequencies of Rearrangements in Thymus and Peripheral T-Lymphocytes

In order to gain further insight into the frequencies of V-J combinations, we set up a precise quantification of rearrangements by real-time genomic quantitative PCR (qPCR). Particular V and J genes were selected as representative of several locations in the TCRAD locus and qPCR was carried out with DNA from thymi (Fig. 4A) and from peripheral blood lymphocytes (PBLs) (Fig. 4B). While the patterns of V-J combinations appear similar among individuals and follow the general rules, some discrete differences in recombination frequencies are detected when comparing the patterns obtained in the thymus and peripheral T-lymphocyte DNA. Several observations emerge from these detailed analyses. Firstly, some V-J combinations (i.e., V1-J56, V1-J53, V40-J10 and V41-J10) are not

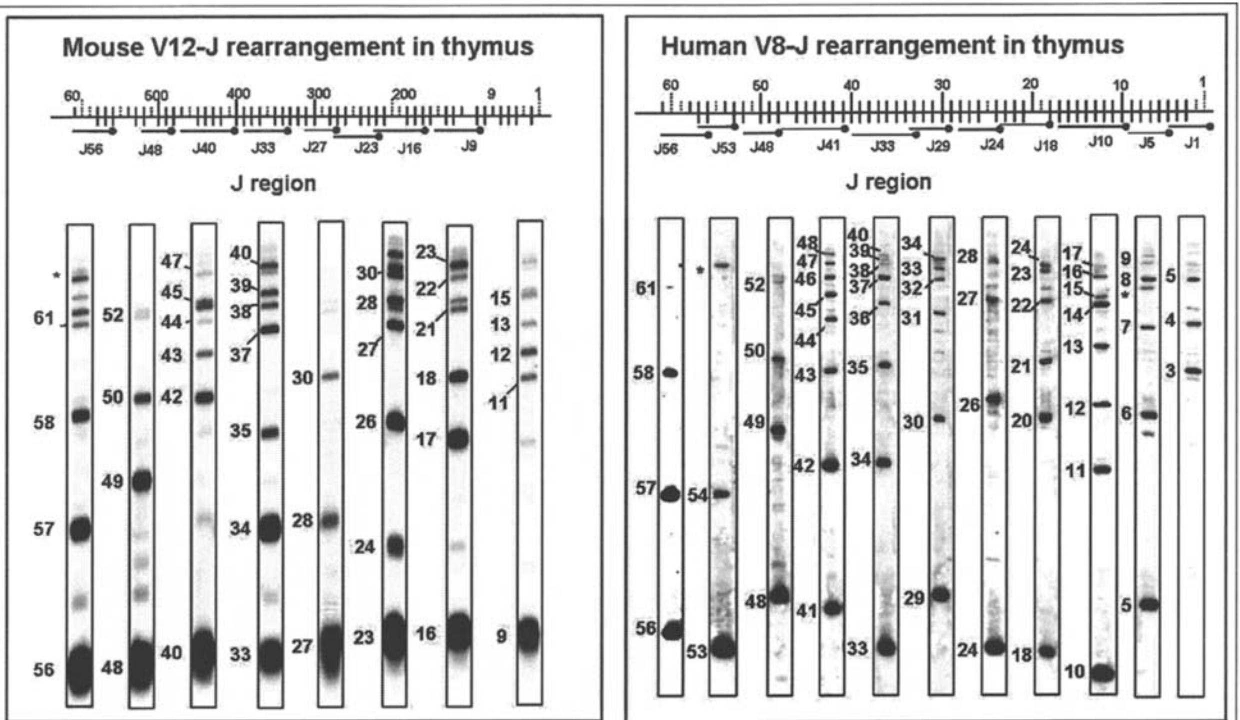


Figure 2. Multiplex PCR assays of V-J rearrangements of the mouse and human TCRAD locus. Genomic DNA was extracted from thymus and amplified using primers situated downstream of different J genes and spread over the J gene region, together with a primer specific for either the mouse V12 (left panel) or human V8 (right panel) gene families. Multiplex PCR was performed with the Expend High Fidelity PCR system (Roche, Meylan, France) as follows: 5 min at 94°C, followed by 26 cycles consisting of 1 min at 94°C, 1 min at 58°C, 6 min at 72°C and finally 10 min at 72°C. The PCRs were in the linear phase of amplification within the 26 cycles. PCR products were migrated on 1.6% agarose gels and probed by Southern hybridization. The probes used to detect the TCRA rearrangements are specific for the respective V genes.

detectable either in the thymus or the PBL, presumably because they are very infrequent. This result confirms the combinatorial pattern described in Figure 3, dependent on the reciprocal position of the V and J genes within the locus. Secondly, some combinations are favored in the periphery with respect to others (for instance V1-J33 can be found at a high frequency in all samples tested). Thirdly, some rearrangements are quantitatively less abundant in the periphery with respect to the thymus. In particular, proximal V-J rearrangements, like V40-J56 or V40-J53, are weakly found (at 6 to 8 cycles of qPCR) in the periphery compared to their high frequency in thymus samples. Several possibilities may account for these differences, including: (1) variation in the number of T-cells between thymus and PBL samples; (2) the contribution of rearrangements occuring on excision circles (these may be more frequently amplified in thymus than in peripheral T-cells in which excision circles have been diluted); (3) the occurrence of secondary rearrangements in the thymus or receptor revision events in the periphery which would replace the most-proximal and accessible V-J rearrangements by joining between more distal V and J genes;[27] (4) positive and negative selection events.[28] Finally, the expansion/contraction of specific rearrangements (i.e., V40-J41, V1-J41, V1-J10) can be identified in certain individuals. Taken together, this analysis demonstrates that, while the recombination pattern is quantitatively similar in thymus samples of several individuals, more heterogeneity of V-J combination is observed in the peripheral T-cell. These observations may indicate the sharing amongst individuals of thymic selection events with similar impact on V-J combination, whereas a divergence amongst individuals in the periphery regarding some V-J combinations could reflect expansions of particular clonotypes induced by immune responses or homeostatic maintenance forces.

The Size of the Mouse and Human TCRα Repertoire

Dependent on the locus position together with the differential expression of V families, preferential V-J recombination leads to a restriction of the potential combinatorial TCR α chain repertoire. By analyzing heterogeneity in CDR3 sequences, the diversity of the human α chain repertoire was estimated at around 0.5×10^6 chains in the blood.[19] However, in this calculation, all the human TCRA V-J combinations were considered as equally likely. The theoretical number of combinations if all 54 V genes could rearrange to each of the 61 J gene segments within the locus is 3294. However, only 46 human V genes and 49 J segments are available for rearrangement. Taking into account (i) that the recombination of proximal V genes including V1.1 to V7 is restricted to the closest half of the J region corresponding to approximately 32 Js; (ii) that the central V genes rearrange with about 45 J gene segments; and (iii) that the distal V genes, (i.e., V31 to V41) do not rearrange with the first 9 Js giving 9 functional V genes rearranging with 40 Js, then the number of possible V-J combinations is less than 2000 ($8V \times 32J + 29V \times 45 + 9V \times 40J$). This suggests that the actual number of combinations corresponds to less than 60% of the estimated 0.5×10^6 total combinatorial possibilities, i.e., 0.3×10^6 TCRα chains. This value is also likely overestimated as it does not take into account the different frequencies of utilization of V and J gene segments within the locus. Concerning mouse, the number of different α chains have been estimated as around 1.2×10^4 in the C57Bl/6 or B10 TCRAD haplotype.[4] It is worth noting that the number of V genes varies from 1 to 3 fold among different haplotypes, for instance the C57Bl/6 haplotype possesses 1/3 less V genes compared to the Balb/c haplotype[12,29] leading to an estimated 0.8×10^4 TCRα chains in the Balb/c haplotype. In addition, multiple rounds of V gene duplications mean that most V families contain between 2 and 10 members, in some cases perhaps differing by only one to three punctual mutations scattered through the V genes.[30] This prevents a precise determination of the number of J segments used by V genes. In the Balb/c TCRA haplotype, (i) the most proximal V genes (V21 to V23) are found rearranged to less than 10 J genes (those between J60 to J48), (ii) the middle V genes use a panel of about 35 Js and finally (iii) the distal V genes (V1 to V3.1) use a panel of less than 30 J segments. Using this information, we estimated a reduction of around 30% in the number of V-J combinations in Balb/c mice compared to the theoritical number of combinations (ref.14 and our unpublished results) yielding an estimated 0.6×10^4 different α chains. Taken together, these findings indicate that whilst remaining large enough to maintain a high functional diversity, limitations of combinatorial diversity reduce the size of the available human and mouse TCR α chain repertoires.

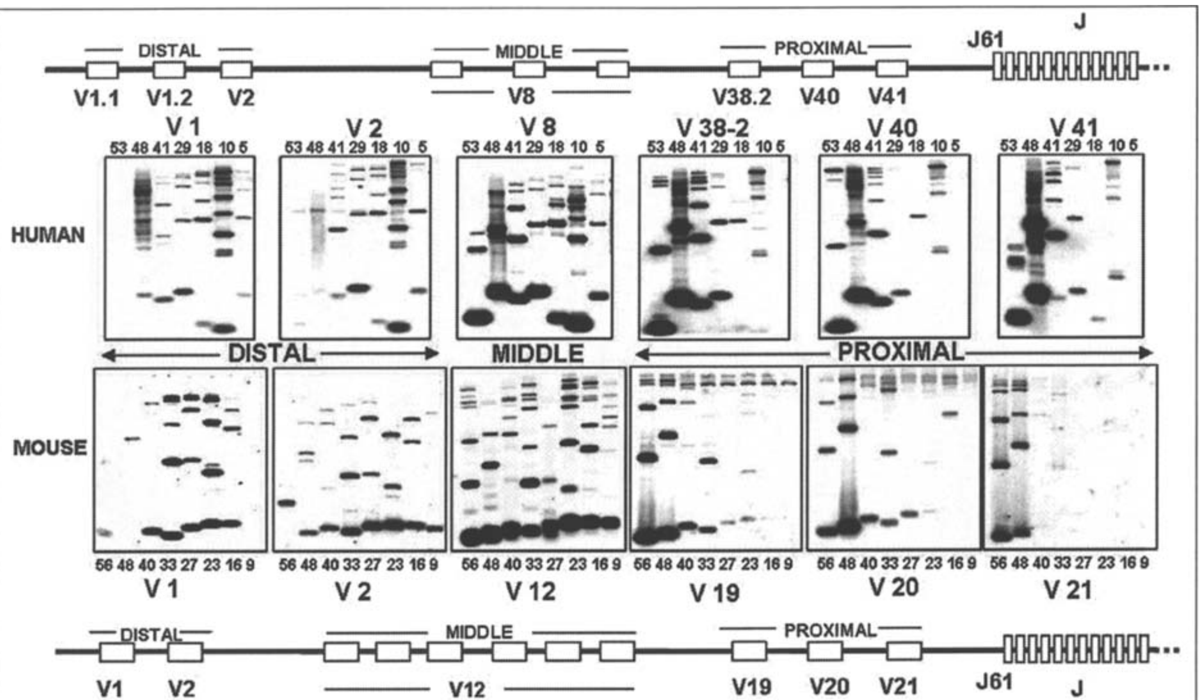


Figure 3. V-J rearrangement depends on V and J localization within the TRAD locus. V-J specific rearrangements were analyzed by multiplex PCR on DNA extracted from total thymus of a 6 day-old human (upper panel) and from total thymus of a 6 week old Balb/c mouse (lower panel). Products of PCR were separated by gel electrophoresis, Southern blotted and hybridized with radioactively labeled probes specific for J genes. Specific V gene primers are indicated on the top or the bottom of each panel. They were used in conjunction with seven representative J primers, as indicated for each lane of the electrophoresis gels. Each band corresponds to a rearrangement event as determined by distance migration. The schematic representation of the TCRAD locus (not drawn to scale) indicates the relative position of each V gene in human (top) and mouse (bottom). For simplicity only the studied V genes and a part of the J gene region are indicated.

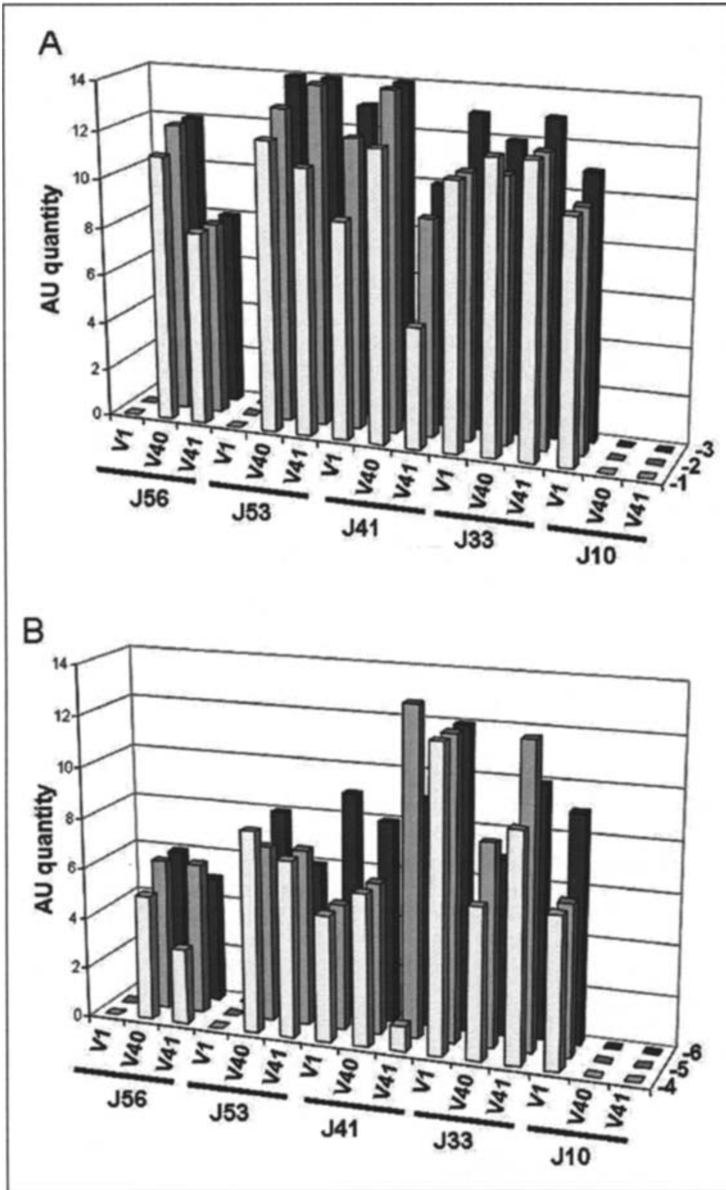


Figure 4. Relative abundance of V-J specific rearrangements among healthy individuals as determined by quantitative genomic PCR analysis. The investigated rearrangements involved V1, 40, 41 and J56, 53, 41, 33, 10 of the human TCRAD locus. The results are expressed in arbitrary units (AU) indicating the differences in cycle numbers at which the products were first detected, therefore reflecting the relative quantities of PCR products for each V-J rearrangement in different individuals. The figure depicts examples of representative results for 3 thymi (A) and 3 PBL (B) DNA samples belonging to 6 healthy individuals aged between 25 and 55 yrs (numbered 1 to 6). Normalization for the DNA content of each sample was performed by amplification of the G3PDH gene. Data are representative of three different experiments.

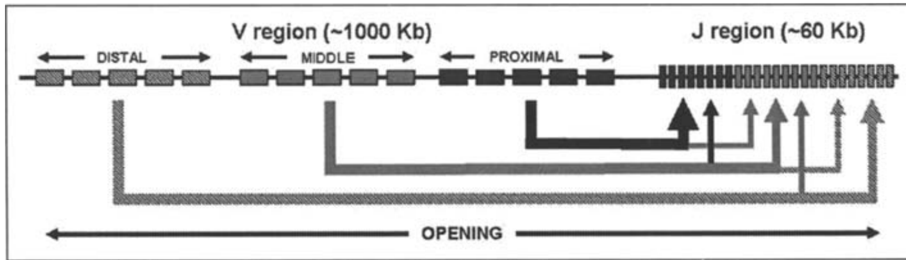


Figure 5. Schematic representation of the V-J combinations in TCRA rearrangements. V and J genes were respectively categorized according to their respective relative locations in the TCRA locus as distal (white dashed boxes), middle (grey boxes) and proximal (black boxes). Combinations of V to J genes are indicated by arrows, with the same color code, where the thickness of the line is indicative of the relative frequencies of V-J combinations. The opening of the TCRA locus to V-J gene rearrangement appears as concentric, from the closest to the most distant V and J genes.

Conclusion

The fact that V and J gene segments combine preferentially according to their position in the TCRA locus suggests a control of rearrangements depending mostly on the strict regulation of chromatin accessibility in both the V and J gene regions (Fig.5). Cis-acting elements, particularly enhancers and promoters, have been proposed as being involved in chromatin remodelling.[31,32] In the murine TCRA locus, accessibility of the J region is controlled by the Eα enhancer located 3′ of the C coding region.[33] In addition, two promoters contribute to the control of Jα rearrangements, namely the T early α (TEA) at the 5′ end of the J region and a second promoter located 15 kb downstream of TEA before the J49 coding region. Both promoters can be activated by Eα.[34,35] The TEA promoter has been shown to spatially regulate J gene utilization[36] and drive noncoding transcription to positively and negatively instruct the activity of downstream J promoters.[37] Interestingly, TEA transcription has been proposed to target V rearrangements to the 5′ end of the J region and consequently determines the rearrangement profile of this region by promoting the activation of proximal J promoters (J58 to J56) while repressing that of more distal J promoters (see chapter by Abarrategui and Krangel). These recent data on the role of TEA transcription on J gene accessibility support the recombination profiles discussed in this report. Whilst we are beginning to gain a better understanding of the mechanisms contributing to the use of J segments, the process of V gene accessibility to rearrangement and the control of their uses remain to be elucidated.

The evaluation of the TCR repertoire is an important measure of the immune competence of an individual. It is assumed that the larger the number of distinct immune T-cells, the more efficient the protection against infectious diseases. Consequently, the size and diversity of the available repertoire are crucial in shaping the immune response to a given antigen. Our studies strongly suggest that although it remains large enough to maintain a high functional diversity, the TCR repertoire of human and mouse α chains is smaller than that predicted by the random rearrangement model. Detailed knowledge about the extent and diversity of the TCR repertoire used in specific immune responses will facilitate the ability to understand the role of the TCR genes in normal and disease states. Whereas clonal populations are hallmarks of malignancy, clonal or oligoclonal populations of T- and B-lymphocytes may also arise in nonmalignant conditions, including normal individuals (responses against some pathogens such as HIV and EBV), elderly patients and patients suffering from autoimmunity or immunodeficiency. Our straightforward experimental approach enables a qualitative and quantitative description of the overall TCRα chain diversity in humans and offers a unique opportunity to characterize and track the repertoire for each individual in healthy and diseased states.

Acknowledgements

The authors wish to thank J Conway, MA Thelu and HL Dougier for their helpful discussions and comments. This work was supported by institutional grants from the Institut National de la Santé et de la Recherche Médicale and from the Commissariat à l'Energie Atomique and partly by specific grants from the Agence Nationale pour la Recherche sur le SIDA et les Hépatites and the Ligue Contre le Cancer. P Fuschiotti was recipient of a fellowship from the Centre National de la Recherche Scientifique.

References

1. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. Nature 1988; 334(6181):395-402.
2. Davis MM, Boniface JJ, Reich Z et al. Ligand recognition by alpha beta T-cell receptors. Annu Rev Immunol 1998; 16:523-544.
3. Cobb RM, Oestreich KJ, Osipovich OA et al. Accessibility control of V(D)J recombination. Adv Immunol 2006; 91:45-109.
4. Cabaniols JP, Fazilleau N, Casrouge A et al. Most alpha/beta T-cell receptor diversity is due to terminal deoxynucleotidyl transferase. J Exp Med 2001; 194(9):1385-1390.
5. Malissen M, Trucy J, Letourneur F et al. A T-cell clone expresses two T-cell receptor alpha genes but uses one alpha beta heterodimer for allorecognition and self MHC-restricted antigen recognition. Cell 1988; 55(1):49-59.
6. Saito T, Sussman JL, Ashwell JD et al. Marked differences in the efficiency of expression of distinct alpha beta T-cell receptor heterodimers. J Immunol 1989; 143(10):3379-3384.
7. Mason D. A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunol Today 1998; 19(9):395-404.
8. Reiser JB, Darnault C, Gregoire C et al. CDR3 loop flexibility contributes to the degeneracy of TCR recognition. Nat Immunol 2003; 4(3):241-247.
9. Thompson SD, Pelkonen J, Hurwitz JL. First T-cell receptor alpha gene rearrangements during T-cell ontogeny skew to the 5′ region of the J alpha locus. J Immunol 1990; 145(7):2347-2352.
10. Roth ME, Holman PO, Kranz DM. Nonrandom use of J alpha gene segments. Influence of V alpha and J alpha gene location. J Immunol 1991; 147(3):1075-1081.
11. Rytkonen MA, Hurwitz JL, Thompson SD et al. Restricted onset of T-cell receptor alpha gene rearrangement in fetal and neonatal thymocytes. Eur J Immunol 1996; 26(8):1892-1896.
12. Jouvin-Marche E, Aude-Garcia C, Candeias S et al. Differential chronology of TCRADV2 gene use by alpha and delta chains of the mouse TCR. Eur J Immunol 1998; 28(3):818-827.
13. Huang C, Kanagawa O. Ordered and coordinated rearrangement of the TCR alpha locus: role of secondary rearrangement in thymic selection. J Immunol 2001; 166(4):2597-2601.
14. Pasqual N, Gallagher M, Aude-Garcia C et al. Quantitative and qualitative changes in V-J alpha rearrangements during mouse thymocytes differentiation: implication for a limited T-cell receptor alpha chain repertoire. J Exp Med 2002; 196(9):1163-1173.
15. Sant'Angelo DB, Waterbury PG, Cohen BE et al. The imprint of intrathymic self-peptides on the mature T-cell receptor repertoire. Immunity 1997; 7(4):517-524.
16. van Meerwijk JP, Marguerat S, Lees RK et al. Quantitative impact of thymic clonal deletion on the T-cell repertoire. J Exp Med 1997; 185(3):377-383.
17. Correia-Neves M, Waltzinger C, Mathis D et al. The shaping of the T-cell repertoire. Immunity 2001; 14(1):21-32.
18. Pannetier C, Even J, Kourilsky P. T-cell repertoire diversity and clonal expansions in normal and clinical samples. Immunol Today 1995; 16(4):176-181.
19. Arstila TP, Casrouge A, Baron V et al. A direct estimate of the human alphabeta T-cell receptor diversity. Science 1999; 286(5441):958-961.
20. Lefranc MP. IMGT, the international immunogenetics database. Nucleic Acids Res 2001; 29(1):207-209.
21. Baum TP, Pasqual N, Thuderoz F et al. IMGT/GeneInfo: enhancing V(D)J recombination database accessibility. Nucleic Acids Res 2004; 32(1):D51-54.
22. Baum TP, Hierle V, Pasqual N et al. IMGT/GeneInfo: T-cell receptor gamma TRG and delta TRD genes in database give access to all TR potential V(D)J recombinations. BMC Bioinformatics 2006; 7(1):224.
23. Koop BF, Rowen L, Wang K et al. The human T-cell receptor TCRAC/TCRDC (C alpha/C delta) region: organization, sequence and evolution of 97.6 kb of DNA. Genomics 1994; 19(3):478-493.
24. Glusman G, Rowen L, Lee I et al. Comparative genomics of the human and mouse T-cell receptor loci. Immunity 2001; 15(3):337-349.

25. Fuschiotti P, Pasqual N, Hierle V et al. Analysis of the TCR alpha-chain rearrangement profile in human T-lymphocytes. Mol Immunol 2007; 44(13):3380-3388.
26. Aude-Garcia C, Gallagher M, Marche PN et al. Preferential ADV-AJ association during recombination in the mouse T-cell receptor alpha/delta locus. Immunogenetics 2001; 52(3-4):224-230.
27. Mostoslavsky R, Alt FW. Receptor revision in T-cells: an open question? Trends Immunol 2004; 25(6):276-279.
28. Bandeira A, Faro J. Quantitative constraints on the scope of negative selection: robustness and weaknesses. Trends Immunol 2003; 24(4):172-173.
29. Jouvin-Marche E, Hue I, Marche PN et al. Genomic organization of the mouse T-cell receptor Valphafamily. EMBO J 1990; 9:2141-2150.
30. Gahery-Segard H, Jouvin-Marche E, Six A et al. Germline genomic structure of the B10. A mouse Tcra-V2 gene subfamily. Immunogenetics 1996; 44(4):298-305.
31. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
32. Mostoslavsky R, Alt FW, Bassing CH. Chromatin dynamics and locus accessibility in the immune system. Nat Immunol 2003; 4(7):603-606.
33. Sleckman BP, Bassing CH, Bardon CG et al. Accessibility control of variable region gene assembly during T-cell development. Immunol Rev 1998; 165:121-130.
34. Villey I, Caillol D, Selz F et al. Defect in rearrangement of the most 5′ TCR-J alpha following targeted deletion of T early alpha (TEA): implications for TCR alpha locus accessibility. Immunity 1996; 5(4):331-342.
35. Hawwari A, Bock C, Krangel MS. Regulation of T-cell receptor alpha gene assembly by a complex hierarchy of germline J alpha promoters. Nat Immunol 2005; 6(5):481-489.
36. Mauvieux L, Villey I, de Villartay JP. T early alpha (TEA) regulates initial TCRVAJA rearrangements and leads to TCRJA coincidence. Eur J Immunol 2001; 31(7):2080-2086.
37. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor alpha recombination. EMBO J 2007; 26(20):4380-4390.

CHAPTER 8

Germline Transcription:
A Key Regulator of Accessibility and Recombination

Iratxe Abarrategui and Michael S. Krangel*

Abstract

The developmental control of V(D)J recombination is imposed at the level of chromatin accessibility of recombination signal sequences (RSSs) to the recombinase machinery. *Cis*-acting transcriptional regulatory elements such as promoters and enhancers play a central role in the control of accessibility in vivo. However, the molecular mechanisms by which these elements influence accessibility are still under investigation. Although accessibility for V(D)J recombination is usually accompanied by germline transcription at antigen receptor loci, the functional significance of this transcription in directing RSS accessibility has been elusive. In this chapter, we review past studies outlining the complex relationship between V(D)J recombination and transcription as well as our current understanding on how chromatin structure is regulated during gene expression. We then summarize recent work that directly addresses the functional role of transcription in V(D)J recombination.

Introduction

V(D)J recombination at antigen receptor loci takes place within the complex nucleoprotein environment of chromatin. An extensive body of literature supports the notion that chromatin-embedded recombination signal sequences (RSSs) must be made accessible to the recombinase for the V(D)J recombination reaction to proceed and that the regulation of RSS accessibility provides an important layer of developmental control to V(D)J recombination in vivo.[1] Studies of antigen receptor loci have implicated promoters and enhancers as developmental regulators of both chromatin structure and V(D)J recombination. However, the detailed mechanisms by which these elements stimulate recombination are not well understood. Enhancers and promoters serve as docking sites for the recruitment of factors that initiate changes in chromatin structure. They also serve as critical regulators of transcription. Studies of antigen receptor loci have demonstrated that unrearranged gene segments typically become transcriptionally active at the developmental stage at which they undergo V(D)J recombination. Nevertheless, whether transcription plays a direct role in providing the recombinase machinery access to RSSs, or is simply an unrelated consequence of locus accessibility, has remained obscure for two decades. Resolution of this issue has required an experimental approach that can discriminate and independently evaluate the individual downstream consequences of enhancer and promoter activity as they relate to the stimulation of V(D)J recombination in vivo. Recent studies have provided important steps forward in this regard and implicate germline transcription as a key developmental regulator of accessibility for V(D)J recombination.

*Corresponding Author: Michael S. Krangel—Department of Immunology, Duke University Medical Center, Durham NC 27710, USA. Email: krang001@mc.duke.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

A Brief History of Germline Transcription and V(D)J Recombination

It was first observed more than 20 years ago by Yancopoulos and Alt that the developmental activation of V_H segment recombination at the *Igh* locus coincided with the appearance of V_H germline transcription.[2] Germline transcripts initiating from promoters associated with V, D and J gene segments have since been documented at all antigen receptor loci and have been shown to coincide developmentally with the onset of V(D)J recombination.[2-5] In addition to these examples of sense transcription across antigen receptor gene segments, recent studies have described a developmental relationship between antisense intergenic transcription across the V_H locus and recombination of V_H gene segments.[6] On the basis of such correlations, germline transcription has long been proposed to play a role in the establishment of an open chromatin configuration that stimulates RSS accessibility.

A linkage between germline transcription and recombination competence was reinforced over the years by a variety of experimental approaches. For example, stable transfection of preB-cells with a recombination substrate containing an exogenous promoter demonstrated that actively transcribed substrates underwent D_H-J_H recombination.[7] Lipopolysaccharide treatment of preB-cells induced both *Igk* locus transcription and Vκ-Jκ rearrangement.[8] Stable transfection of preB-cells with recombination substrates showed that an enhancer promoted both recombination and transcription.[9] Likewise, the introduction of a strong promoter into the *Igl* locus by homologous recombination caused a dramatic increase in both Jλ germline transcription and Vλ-Jλ recombination.[10]

Several transcription factors have also been shown to coordinately regulate both trancription and V(D)J recombination. Overexpression of E2A in recombinase-expressing nonlymphoid cells induced both germline transcription and recombination of *Igk*, *Tcrg* and *Tcrd* gene segments.[11,12] Mice deficient for the transcription factor OcaB displayed defective transcription and recombination of a subset of Vκ genes.[13] In addition, Stat5 was shown to be required for transcription and recombination of distal V_H segments and Jγ segments,[14,15] in response to IL-7 receptor signaling and for transcription and recombination of the Vγ5 gene segment in response to IL-15 receptor signaling.[16] Consistent with all of the above, gene targeting experiments have shown deletion of enhancers and promoters at endogenous loci to inhibit both transcription and V(D)J recombination of linked gene segments.[1] However, none of the above studies had the power to critically test a causal relationship between transcription and V(D)J recombination.

In contrast, several other studies have indicated that V(D)J recombination and germline transcription are not invariably linked. In some instances, transcription through gene segments was shown to be insufficient to promote recombinase activity. For example, distal V_H gene segments are transcribed at high levels in *Pax5*$^{-/-}$ pro-B-cells but fail to undergo recombination.[17] However, these transcribed V_H segments might retain a permissive chromatin configuration in the absence of Pax5, but might fail to rearrange due to additional Pax5 functions that are needed for recombination. For example, Pax5 has been shown to regulate an *Igh* locus conformational change that is required to bring distal V_H and DJ_H segments into proximity for V_H to DJ_H recombination and to recruit RAG proteins to V_H segments.[18,19] Several studies using versions of a transgenic TCRβ minilocus recombination substrate have also provided examples of transcription in the absence of recombination. In one case, Eβ and Eμ were shown to stimulate substrate Dβ to Jβ but not Vβ to DβJβ recombination in B-cells, even though these enhancers could promote germline transcription of Vβ and DβJβ segments in those cells.[20] In other instances, minimal forms of Eβ or Eμ that lacked binding sites for specific nuclear factors efficiently stimulated minilocus transcription but could not support recombination.[21,22] These results suggest that there may be enhancer functions that promote recombination independent of enhancer effects on transcription, but do not rule out a role for enhancer-directed transcription.

Studies of the *Tcrb* locus have also described circumstances in which active transcription is not predictive of recombination. For example, germline transcription of Vβ8.2 occurs on both alleles in all $CD4^-CD8^-$ double negative (DN) thymocytes even though Vβ8.2 usually rearranges on only a single allele and in only a fraction of these cells.[23] Ectopic introduction of Eα downstream

of Vβ12 greatly enhanced the transcription of this segment in DP thymocytes but did not induce Vβ12-DβJβ recombination.[24] Similarly, a large *Tcrb* locus deletion that placed Vβ segments under the influence of Eβ stimulated high level Vβ transcription but not recombination in DP thymocytes.[25] However, the failure to rearrange in these examples may be explained not by any lack of Vβ segment accessibility but by inappropriate nuclear localization or locus conformation. Alternatively, there may be constraints imposed by unknown factors that might promote or inhibit Vβ to DJβ recombination at the appropriate developmental stage.

In other instances, recombination has been documented to occur in the absence of germline transcription. A study using isolated nuclei from RAG deficient cells showed that endogenous RSSs could be cleaved by the addition of RAG proteins in vitro in the absence of ongoing transcription.[26] However, chromatin modifications introduced by transcription prior to the isolation of nuclei could have been sufficient to provide gene segment accessibility during subsequent in vitro cleavage reactions. Another study identified endogenous V_H segment rearrangements in pro-B-cells that did not detectably transcribe those V_H segments.[27] In this system, rearrangement was induced by transfection of RAG-deficient pro-B-cell clones with RAG expression plasmids. However, the transcription status of V_H segments at the time of recombination could not be analyzed, leaving open the possibility that V_H transcription did occur in the small fraction of cells that underwent rearrangement during the three-week culture period following RAG transfection.

Several studies have shown that localized chromatin remodeling at promoters is sufficient to stimulate recombination at adjacent gene segments in the absence of read-through transcription. In one example, an RSS was tightly associated with the inducible mouse mammary tumor virus long terminal repeat.[28] When propagated as a chromatinized episomal substrate, nucleosome organization at the promoter precluded protein access. However, the mobilization of promoter nucleosomes by treatment with dexamethasone was found to promote recombination in the absence of measurable transcription. At the endogenous *Tcrb* locus, accessibility at Dβ1 requires the concerted action of Eβ and a promoter tightly associated with this segment, PDβ1.[29-32] A physical interaction between Eβ and PDβ1 is required to deliver the SWI/SNF chromatin-remodeling complex to PDβ1, resulting in decreased nucleosome occupancy at Dβ1.[31,32] A series of experiments making use of a TCRβ minilocus have argued that Dβ1 accessibility depends on local targeting of SWI/SNF by PDβ1 but can occur independent of PDβ1-derived transcription. For example, minilocus recombination requires that PDβ1 is situated immediately adjacent to the Dβ1 RSS, but can be supported by a version of PDβ1 that does not stimulate transcription through the Dβ and Jβ segments.[33] Moreover, PDβ1 function could be substituted by controlled targeting of the catalytic subunit of SWI/SNF to Dβ1.[34] These studies argue persuasively that transcription is not an absolute requirement for accessibility, particularly when a promoter and RSS are tightly associated. However, these studies do not discount the possibility that transcription could contribute substantially to accessibility at endogenous antigen receptor loci in vivo.

Disruption of Chromatin by Transcription

The genetic material is packed in the eukaryotic nucleus in a highly organized fashion.[35] The first level of compaction is achieved by wrapping 146 bp of DNA around the histone octamer (composed of two copies each of histone H2A, H2B, H3 and H4) to form the nucleosome particle.[36] A linear array of nucleosomes, with 20-60 bp of internucleosomal linker DNA, forms the 10 nm fiber that has the appearance of 'beads on a string' when viewed under an electron microscope. Formation of the more compact 30 nm fiber depends on the binding of histone H1 to linker DNA and on the establishment of internucleosomal interactions. However, the mechanisms that govern compaction of the 30 nm fiber into higher order structures, ultimately resulting in the assembly of chromosomes, remain elusive.[37] The highly compact chromatin organization inhibits access of proteins to the underlying DNA, thereby imposing an obstacle to transcription. Eukaryotic cells use a variety of strategies to dynamically modulate chromatin structure to achieve regulated gene expression.

All four histones are subjected to a variety of posttranslational modifications that include acetylation, methylation, phosphorylation and ubiquitylation.[38] These modifications are targeted

both to the extended amino terminal tails of histones that project away from the nucleosome surface and to the globular domains of histones. Specific patterns of histone modifications correlate well with the activation status of a gene. For example, active genes typically display high levels of histone H3 and H4 acetylation and histone H3 lysine 4 (H3 K4) methylation, whereas repressed genes are typically enriched for histone H3 K9 and K27 methylation. Histone modifications are highly dynamic. For instance, the introduction of histone acetylation by histone acetyltransferases (HATs) is reversed by histone deacetylases (HDACs),[39] and the introduction of histone methylation by histone methyltranferases (HMTs) is reversed by histone demethylases.[40] Moreover, the introduction of histone marks can be influenced by chromatin context, in the sense that prior modifications to neighboring amino acids can either promote or inhibit the introduction of a subsequent modification.[41,42] The function of these modifications is twofold: on one hand, they loosen or decompact chromatin structure,[43] and on the other hand, they recruit multiprotein effector complexes that directly regulate chromatin structure and function.[44-47] These effectors are targeted to chromatin based on the properties of the various chromatin binding modular protein domains that they contain. For example, bromodomain-containing proteins bind acetylated histones, whereas chromodomain and PHD finger motif-containing proteins recognize histones methylated at different lysine residues.[39,42] Proteins or protein complexes with combinations of chromatin binding domains may be preferentially recruited to nucleosomes displaying more complex patterns of histone modification.[48]

Chromatin structure can also be modulated through the activity of ATP-dependent chromatin-remodeling complexes, which use the energy derived from ATP hydrolysis to disrupt histone-DNA contacts.[49,50] These remodeling complexes can be categorized into three major families (SWI2/SNF2, ISWI and Mi2/CHD) based on the structures of their ATPase catalytic subunit. The function and the recruitment of these complexes are determined by the domain structures of their ATPase subunit and of additional subunits within the complex. For example, the BRM and BRG1 catalytic subunits of the SWI2/SNF2 complex contain a bromodomain that targets the complex to acetylated histones to positively or negatively regulate gene expression. The SANT and SLIDE domains of ISWI recognize histone tails and linker DNA, respectively. ISWI complexes play a central role in the ordering and spacing of nucleosomes to promote gene repression. However, when ISWI is part of the PHD finger motif-containing NURF complex, its recruitment to H3 K4 trimethylated histones is required for proper gene activation.[47] The ATPase subunit Mi-2 contains a chromodomain that directs binding to nucleosomes; Mi-2 complexes generally contain HDAC subunits that contribute to transcriptional repression.

The exact mechanisms by which these complexes induce nucleosome remodeling are not fully elucidated.[50] Several studies have shown that changes in nucleosome structure can be promoted through an initial DNA translocation within the nucleosome which is then propagated as a DNA bulge around the histone octamer. This leads to nucleosome sliding with respect to the DNA sequence and, in some cases, to nucleosome disassembly.[51] Ultimately, nucleosome remodeling causes an increase in accessibility of the underlying DNA to transcription factors.

Gene expression is initiated by the binding of transcriptional activators to promoter regions.[41,52] This is then followed by the sequential recruitment of histone-modifying enzymes and chromatin-remodeling complexes to the promoter. These chromatin regulators act cooperatively to disrupt promoter nucleosomes and to allow the formation of a stable preinitiation complex at the promoter. For example, acetylation of promoter nucleosomes can help to recruit the SWI/SNF chromatin-remodeling complex, resulting in nucleosome displacement or disassembly prior to gene activation.[53] Promoter-engaged RNA pol II complexes subsequently transit through chromatin to mediate transcriptional elongation.

Transition into the elongation phase of transcription (clearance of RNA pol II from the promoter) requires phosphorylation of the RNA pol II carboxy-terminal domain (CTD).[54] The RNA pol II CTD represents a platform for the recruitment of histone modifying complexes and elongation factors that allow the polymerase to efficiently transcribe through chromatin. For instance, the PAF complex is required for the recruitment of Set1[55] (responsible for H3 K4 methylation)

and Rad6/Bre[56] (responsible for H2B monoubiquitination) to elongating RNA pol II, whereas Set2[57] (responsible for H3 K36 methylation) binds directly to the phosphorylated RNA pol II CTD. Among the chromatin modifications introduced by these complexes, H3 K4 methylation recruits additional chromatin regulators that in turn remodel nucleosomes within transcribed regions.[45-47] Histone acetylation and monoubiquitination stimulate transcription by promoting nucleosome disassembly ahead of the polymerase.[58,59] Finally, H3 K36 methylation and histone deacetylation reestablish the chromatin structure by repositioning nucleosomes after RNA pol II passage.[60,61] Several histone chaperones that travel with RNA pol II also mediate histone H2A/H2B and H3/H4 eviction ahead of the polymerase and the subsequent reassembly of nucleosomes in its wake.[62,63] In addition, chromatin-remodeling complexes are targeted to transcribed coding regions to facilitate the passage of RNA pol II through the nucleosomal barrier.[64,65] Thus, chromatin disruption is an intrinsic feature of the process of transcription.

Regulation of V(D)J Recombination by Transcription

We have recently addressed the role of transcriptional elongation in the control of accessibility for recombination at the mouse *Tcra* locus. The Jα region near the 3' end of the *Tcra* locus contains 61 Jα segments that span 65 kilobases (Fig. 1A).[66] All Vα to Jα recombination events depend on the *Tcra* enhancer (Eα) located at the extreme 3' end of the locus.[67] Primary Vα to Jα rearrangements are targeted to the most 5' Jα segments by the activity of two germline promoters that are activated by Eα.[68,69] These promoters, TEA at the 5' end of the Jα array and the Jα49 promoter 15 kb downstream, drive the expression of germline transcripts that extend across the entire Jα-Cα region. Subsequently, secondary Vα to Jα rearrangement events replace the primary rearrangements resulting in the use of progressively more 5' Vα and 3' Jα segments.[70,71] Secondary rearrangements are thought to be targeted to Jα segments downstream of a primary VαJα rearrangement by the promoter of the rearranged Vα gene segment.[72]

Studies performed on TEA-deleted mice have shown that TEA controls rearrangements and chromatin structure across a 12 kb region encompassing the most 5' Jα segments (Jα61-Jα52) (Fig. 1B).[68,69,73] Within the TEA-dependent accessibility region, Jα58, Jα57 and Jα56 segments are associated with promoters whose activation depends on TEA.[69] However, Jα61, Jα53 and Jα52 lack their own promoters and are located at a distance from the nearest upstream promoters. Jα61 is situated 1.7 kb downstream of TEA and Jα53 and Jα52 are located 3.4 and 6.1 kb, respectively, from the nearest upstream promoter at Jα56. We wondered whether accessibility of these promoter-distal Jα segments might depend on transcription from upstream promoters.

To critically test the regulatory function of transcription in Jα segment recombination, we used homologous recombination to introduce a strong transcription terminator cassette downstream of Jα56 (Fig. 1C).[74] The terminator is composed of a set of four polyadenylation sites followed by an array of twelve bacterial *lac* operators. The *lac* operators are thought to function as strong pause sites for RNA pol II, which would increase the efficiency of both polyadenylation and termination. The introduced transcription terminator was shown by reverse transcriptase polymerase chain reaction and by nuclear run-on to impose an effective block to RNA pol II passage through the Jα53-Jα52 region in vivo, virtually eliminating germline transcription across these segments. Moreover, the transcriptional block caused an 87% reduction in recombination at Jα53 and a 67% reduction at Jα52. As expected, recombination of upstream and downstream Jα segments was unaffected, due to the presence of additional promoters associated with these gene segments. Transcriptional blockade also led to very specific alterations in chromatin structure at Jα53 and Jα52. Acetylation of histones H3 and H4 was unaffected. However, dimethylation and trimethylation of histone H3 K4 were both reduced at these Jα segments. These findings provided the first direct experimental proof that transcriptional elongation across promoter-distal gene segments is required for normal rates of recombination in vivo and suggested a potential link between transcription-dependent H3 K4 methylation and accessibility to the recombinase.

The transcription terminator approach also provided insights into the mechanisms that regulate germline Jα promoter activity. The Jα region contains a series of cryptic promoters located in the

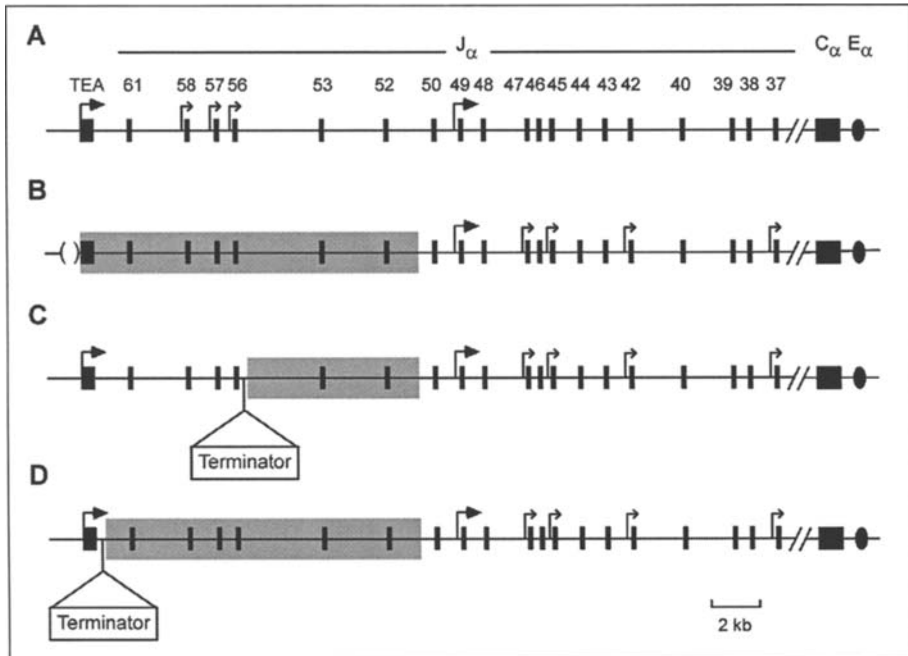


Figure 1. Regulation of *Tcra* locus Jα segment recombination by transcription. A) Wild-type Jα-Cα region, identifying active promoters (bent arrows) and enhancer Eα (filled oval). B) TEA promoter-deleted Jα-Cα region.[73,75] Shading identifies region of reduced transcription, reduced histone acetylation and methylation and reduced recombination. C) Terminator introduction downstream of Jα56.[74] Shading identifies region of reduced transcription, reduced histone methylation and reduced recombination. D) Terminator introduction downstream of TEA.[75] Shading identifies region of reduced transcription, reduced histone acetylation and methylation and reduced recombination.

central portion of the array (Jα47-Jα37) that become activated when the TEA promoter is deleted (Fig. 1B).[69] Like TEA deletion, transcriptional blockade also caused the de-repression of these cryptic Jα promoters, showing that their activity is normally suppressed through a transcriptional interference mechanism (Fig. 1C).[74] Suppression of these promoters by transcriptional interference is likely to be important because when these promoters are active they target rearrangement events to the central Jα segments and lead to disordered usage of Jα segments.[69]

More recently, we extended this analysis by introducing the transcription terminator downstream of the TEA promoter (Fig. 1D).[75] In this location the terminator almost completely eliminated recombination of Jα61 through Jα52, mimicking the phenotype of TEA-deleted mice. This occurs in part because TEA transcription is required for the activity of promoters associated with Jα58, Jα57 and Jα56, even as it suppresses the activity of central Jα promoters.

In conclusion, transcriptional elongation can stimulate accessibility for recombination in two different ways. On one hand, transcription can directly provide long-range accessibility at promoter-distal RSSs, probably through alterations in chromatin structure associated with the process of transcriptional elongation. On the other hand, transcription can regulate the activity of additional promoters which themselves can provide accessibility to nearby RSSs.

We think that transcription is likely to contribute to the regulation of RSS accessibility at other immunoglobulin and T-cell receptor loci, but that the extent of contribution may depend on the proximity of the RSS to the nearest promoter. Dβ1 RSS accessibility may occur independent of

transcription because the Dβ1 RSS is close enough to the promoter that it can be influenced by chromatin remodeling complexes that are targeted to the promoter. V segment promoters are generally located three to four nucleosomes away from the V segment RSS. Therefore, accessibility at these RSS might be influenced by both transcription-dependent and promoter-targeted chromatin remodeling. At the *Tcrb* locus, Jβ1 segments are situated 1-2 kb from PDβ1. Analysis of PDβ1-deleted mice indicated that Jβ1.6 accessibility depends at least in part on PDβ1,[32] and we suggest that this is likely to be a consequence of transcription from PDβ1. Long antisense intergenic *Igh* transcripts may similarly regulate chromatin remodeling events that influence accessibility across the V_H and D_HJ_H regions.[6,76]

Future Directions

The studies outlined above fall short of clarifying the precise mechanisms by which transcription can stimulate accessibility for V(D)J recombination. It will be important to address these mechanistic issues in future studies. Specific histone modifications introduced during transcriptional elongation are likely to stimulate accessibility in a variety of ways. For example, we noted a dramatic suppression of H3 K4 trimethylation at Jα53 and Jα52 as a consequence of transcriptional blockade. H3 K4 trimethylation can recruit PHD finger motif-containing proteins to active genes.[42]. Since this domain is present in several histone-modifying and chromatin remodeling complexes, the recruitment of these additional activities onto chromatin might in turn be important for accessibility.[46,47,59] It is also notable that the C-terminus of RAG2 contains a PHD finger motif that is important for recombinase activity and that mediates interactions with histones.[77,78] Hence there could be a direct mechanistic link between transcription-dependent H3 K4 trimethylation of nucleosomes and recombinase recruitment. H3 and H4 acetylation was not suppressed by blockade of transcription at Jα56, indicating that histone acetylation is not sufficient for accessibility. However, acetylation was suppressed when transcription was blocked across the entire 5' region by terminator insertion downstream of TEA.[75] Transcription-dependent histone acetylation might influence accessibility by promoting general chromatin opening, since histone acetylation is known to reduce the compaction of nucleosome arrays.[43] Several histone modifications introduced during transcription, including acetylation and H2B monoubiquitylation, are thought to function together with histone chaperones to promote the transient disassembly of nucleosomes (for example, eviction of H2A/H2B dimers or of H3 and H4) that is associated with RNA Pol II passage.[58,79] Transient nucleosome disassembly could also play a role in RAG accessibility, particularly if RAG itself could associate with RNA pol II and could be delivered at the appropriate time. Regardless, transient nucleosome disassembly might allow access of other relevant factors to their cognate DNA sequences. For example, Pax5 binding to V_H segments can mediate the recruitment of RAG to V_H segments and promote V_H to D_HJ_H recombination.[19] Transcription-dependent chromatin remodeling might directly stimulate Pax5 access to V_H segments and thus indirectly promote RAG recruitment. More than likely, transcription impacts recombination in several ways. Thus although at one level our studies close a longstanding debate in the accessibility field, they also raise many new questions and suggest avenues for additional research that may ultimately contribute to a detailed mechanistic understanding of accessibility.

Acknowledgements

The authors acknowledge support of the National Institutes of Health.

References

1. Cobb RM, Oestreich KJ, Osipovich OA et al. Accessibility control of V(D)J recombination. Adv Immunol 2006; 91:45-109.
2. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. Cell 1985; 40:271-281.
3. Fondell JD, Marcu KB. Transcription of germ line Vα segments correlates with ongoing T-cell receptor α-chain rearrangement. Mol Cell Biol 1992; 12:1480-1489.
4. Goldman JP, Spencer DM, Raulet DH. Ordered rearrangement of variable region genes of the T-cell receptor γ locus correlates with transcription of the unrearranged genes. J Exp Med 1993; 177:729-739.

5. Schlissel MS, Corcoran LM, Baltimore D. Virus-transformed preB-cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. J Exp Med 1991; 173:711-20.
6. Bolland DJ, Wood AL, Johnston CM et al. Antisense intergenic transcription in V(D)J recombination. Nat Immunol 2004; 5:630-637.
7. Blackwell TK, Moore MW, Yancopoulos GD et al. Recombination between immunoglobulin variable region gene segments is enhanced by transcription. Nature 1986; 324:585-589.
8. Schlissel MS, Baltimore D. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. Cell 1989; 58:1001-1007.
9. Oltz EM, Alt FW, Lin WC et al. A V(D)J recombinase-inducible B-cell line: role of transcriptional enhancer elements in directing V(D)J recombination. Mol Cell Biol 1993; 13:6223-6230.
10. Sun T, Storb U. Insertion of phosphoglycerine kinase (PGK)-neo 5' of Jλ1 dramatically enhances VJλ1 rearrangement. J Exp Med 2001; 193:699-712.
11. Romanow WJ, Langerak AW, Goebel P et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Mol Cell 2000; 5:343-353.
12. Ghosh JK, Romanow WJ, Murre C. Induction of a diverse T-cell receptor ψ/δ repertoire by the helix-loop-helix proteins E2A and HEB in nonlymphoid cells. J Exp Med 2001; 193:769-775.
13. Casellas R, Jankovic M, Meyer G et al. OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin kappa genes. Cell 2002; 110:575-585.
14. Bertolino E, Reddy K, Medina KL et al. Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. Nat Immunol 2005; 6:836-843.
15. Ye SK, Agata Y, Lee HC et al. The IL-7 receptor controls the accessibility of the TCRγ locus by Stat5 and histone acetylation. Immunity 2001; 15:813-823.
16. Zhao H, Nguyen H, Kang J. Interleukin 15 controls the generation of the restricted T-cell receptor repertoire of γδ intestinal intraepithelial lymphocytes. Nat Immunol 2005; 6:1263-1271.
17. Hesslein DG, Pflugh DL, Chowdhury D et al. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev 2003; 17:37-42.
18. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18:411-422.
19. Zhang Z, Espinoza CR, Yu Z et al. Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V_H-to-DJ_H rearrangement of immunoglobulin genes. Nat Immunol 2006; 7:616-624.
20. Okada A, Mendelsohn M, Alt F. Differential activation of transcription versus recombination of transgenic T-cell receptor β variable region gene segments in B and T lineage cells. J Exp Med 1994; 180:261-272.
21. Fernex C, Capone M, Ferrier P. The V(D)J recombinational and transcriptional activities of the immunoglobulin heavy-chain intronic enhancer can be mediated through distinct protein-binding sites in a transgenic substrate. Mol Cell Biol 1995; 15:3217-3226.
22. Tripathi RK, Mathieu N, Spicuglia S et al. Definition of a T-cell receptor β gene core enhancer of V(D) J recombination by transgenic mapping. Mol Cell Biol 2000; 20:42-53.
23. Jia J, Kondo M, Zhuang Y. Germline transcription from T-cell receptor Vβ gene is uncoupled from allelic exclusion. EMBO J 2007; 26:2387-2399.
24. Jackson A, Kondilis HD, Khor B et al. Regulation of T-cell receptor β allelic exclusion at a level beyond accessibility. Nat Immunol 2005; 6:189-197.
25. Senoo M, Wang L, Suzuki D et al. Increase of TCR Vβ accessibility within Eβ regulatory region influences its recombination frequency but not allelic exclusion. J Immunol 2003; 171:829-835.
26. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85:887-897.
27. Angelin-Duclos C, Calame K. Evidence that immunoglobulin V_H-DJ recombination does not require germ line transcription of the recombining variable gene segment. Mol Cell Biol 1998; 18:6253-6264.
28. Cherry SR, Baltimore D. Chromatin remodeling directly activates V(D)J recombination. Proc Natl Acad Sci USA 1999; 96:10788-10793.
29. Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the Dβ1 gene segment at the TCRβ locus by a germline promoter. Immunity 1999; 10:313-322.
30. Mathieu N, Hempel WM, Spicuglia S et al. Chromatin remodeling by the T-cell receptor (TCR)-β gene enhancer during early T-cell development: Implications for the control of TCR-β locus recombination. J Exp Med 2000; 192:625-636.
31. Spicuglia S, Kumar S, Yeh JH et al. Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. Mol Cell 2002; 10:1479-1487.
32. Oestreich KJ, Cobb RM, Pierce S et al. Regulation of TCRβ gene assembly by a promoter/enhancer holocomplex. Immunity 2006; 24:381-391.

33. Sikes ML, Meade A, Tripathi R et al. Regulation of V(D)J recombination: a dominant role for promoter positioning in gene segment accessibility. Proc Natl Acad Sci USA 2002; 99:12309-12314.
34. Osipovich O, Cobb RM, Oestreich KJ et al. Essential function for SWI-SNF chromatin remodeling complexes in the promoter-directed assembly of Tcrb genes. Nat Immunol 2007; 8:809-816.
35. Misteli T. Beyond the sequence: cellular organization of genome function. Cell 2007; 128:787-800.
36. Khorasanizadeh S. The nucleosome: from genomic organization to genomic regulation. Cell 2004; 116:259-272.
37. Tremethick DJ. Higher-order structures of chromatin: the elusive 30 nm fiber. Cell 2007; 128:651-654.
38. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128:693-705.
39. Lee KK, Workman JL. Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol 2007; 8:284-295.
40. Shi Y, Whetstine JR. Dynamic regulation of histone lysine methylation by demethylases. Mol Cell 2007; 25:1-14.
41. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell 2007; 128:707-719.
42. Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 2007; 25:15-30.
43. Shogren-Knaak M, Ishii H, Sun JM et al. Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science 2006; 311:844-847.
44. Lachner M, O'Carroll D, Rea S et al. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 2001; 410:116-120.
45. Shi X, Hong T, Walter KL et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature 2006; 442:96-99.
46. Taverna SD, Ilin S, Rogers RS et al. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol Cell 2006; 24:785-796.
47. Wysocka J, Swigut T, Xiao H et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 2006; 442:86-90.
48. Li B, Gogol M, Carey M et al. Combined action of PHD and Chromodomains directs the Rpd3S HDAC to transcribed chromatin. Science 2007; 316:1050-1054.
49. Saha A, Wittmeyer J, Cairns BR. Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol 2006; 7:437-447.
50. Smith CL, Peterson CL. ATP-dependent chromatin remodeling. Curr Top Dev Biol 2005; 65:115-148.
51. Bruno M, Flaus A, Stockdale C et al. Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities. Mol Cell 2003; 12:1599-1606.
52. Workman JL. Nucleosome displacement in transcription. Genes Dev 2006; 20:2009-2017.
53. Hassan AH, Neely KE, Workman JL. Histone acetyltransferase complexes stabilize SWI/SNF binding to promoter nucleosomes. Cell 2001; 104:817-827.
54. Sims RJ 3rd, Belotserkovskaya R, Reinberg D. Elongation by RNA polymerase II: the short and long of it. Genes Dev 2004; 18:2437-2468.
55. Krogan NJ, Dover J, Wood A et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol Cell 2003; 11:721-729.
56. Wood A, Schneider J, Dover J et al. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem 2003; 278:34739-34742.
57. Xiao T, Hall H, Kizer KO et al. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev 2003; 17:654-663.
58. Govind CK, Zhang F, Qiu H et al. Gcn5 promotes acetylation, eviction and methylation of nucleosomes in transcribed coding regions. Mol Cell 2007; 25:31-42.
59. Pavri R, Zhu B, Li G et al. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell 2006; 125:703-717.
60. Carrozza MJ, Li B, Florens L et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 2005; 123:581-592.
61. Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell 2005; 20:971-978.
62. Belotserkovskaya R, Oh S, Bondarenko VA et al. FACT facilitates transcription-dependent nucleosome alteration. Science 2003; 301:1090-1093.
63. Schwabish MA, Struhl K. Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. Mol Cell 2006; 22:415-422.

64. Carey M, Li B, Workman JL. RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. Mol Cell 2006; 24:481-487.
65. Corey LL, Weirich CS, Benjamin IJ et al. Localized recruitment of a chromatin-remodeling activity by an activator in vivo drives transcriptional elongation. Genes Dev 2003; 17:1392-1401.
66. Krangel MS, Carabana J, Abarrategui I et al. Enforcing order within a complex locus: current perspectives on the control of V(D)J recombination at the murine T-cell receptor α/δ locus. Immunol Rev 2004; 200:224-232.
67. Sleckman BP, Bardon CG, Ferrini R et al. Function of the TCRα enhancer in αβ and γδ T-cells. Immunity 1997; 7:505-515.
68. Villey I, Caillol D, Selz F et al. Defect in rearrangement of the most 5' TCR-Jα following targeted deletion of T early alpha (TEA): implications for TCRα locus accessibility. Immunity 1996; 5:331-342.
69. Hawwari A, Bock C, Krangel MS. Regulation of T-cell receptor-α gene assembly by a complex hierarchy of germline Jα promoters. Nat Immunol 2005; 6:481-489.
70. Buch T, Rieux-Laucat F, Forster I et al. Failure of HY-specific thymocytes to escape negative selection by receptor editing. Immunity 2002; 16:707-718.
71. Wang F, Huang CY, Kanagawa O. Rapid deletion of rearranged T-cell antigen receptor (TCR) Vα-Jα segment by secondary rearrangement in the thymus: role of continuous rearrangement of TCRα chain gene and positive selection in the T-cell repertoire formation. Proc Natl Acad Sci USA 1998; 95:11834-11839.
72. Hawwari A, Krangel MS. Role for rearranged variable gene segments in directing secondary T-cell receptor-α recombination. Proc Natl Acad Sci USA 2007; 104:903-907.
73. Mauvieux L, Villey I, de Villartay J-P. TEA regulates local TCR-Jα accessibility through histone acetylation. Eur J Immunol 2003; 33:2216-2222.
74. Abarrategui I, Krangel MS. Regulation of T-cell receptor-α gene recombination by transcription. Nat Immunol 2006; 7:1109-1115.
75. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor α recombination. EMBO J 2007; 26: 4380-4390.
76. Bolland DJ, Wood AL, Afshar R et al. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Eμ. Mol Cell Biol 2007; 27:5523-5533.
77. Elkin SK, Ivanov D, Ewalt M et al. A PHD finger motif in the C terminus of RAG2 modulates recombination activity. J Biol Chem 2005; 280:28701-28710.
78. West KL, Singha NC, De Ioannes P et al. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. Immunity 2005; 23:203-212.
79. Tanny JC, Erdjument-Bromage H, Tempst P et al. Ubiquitylation of histone H2B controls RNA polymerase II transcription elongation independently of histone H3 methylation. Genes Dev 2007; 21:835-847.

Chapter 9

Dynamic Regulation of Antigen Receptor Gene Assembly

Lance R. Thomas, Robin Milley Cobb and Eugene M. Oltz*

Abstract

A hallmark feature of adaptive immunity is the production of lymphocytes bearing an enormous repertoire of receptors for foreign antigens. This repertoire is generated early in B and T-cell development by the process of V(D)J recombination, which randomly assembles functional immunoglobulin (Ig) and T-cell receptor (TCR) genes from large arrays of DNA segments. Precursor lymphocytes must target then retarget a single V(D)J recombinase enzyme to distinct regions within antigen receptor loci to guide lymphocyte development and to ensure that each mature B and T-cell expresses only a single antigen receptor specificity. Proper targeting of V(D)J recombinase is also essential to avoid chromosomal aberrations that result in lymphoid malignancies. Early studies suggested that changes in the specificity of V(D)J recombination are achieved by differentially opening or closing chromatin associated with Ig and TCR gene segments at the proper developmental time point. This accessibility model has been extended significantly in recent years and it has become clear that control mechanisms governing antigen receptor gene assembly are multifaceted and vary from locus to locus. In this chapter we review how genetic and epigenetic mechanisms as well as widespread changes in chromosomal conformation synergize to orchestrate the diversification of genes encoding B and T-cell antigen receptors.

Introduction

One triumph of vertebrate evolution is the development of an adaptive immune system, which recognizes and eliminates a continually changing spectrum of pathogens. To accomplish this feat, mammals have evolved a "brute-force" approach to adaptive immunity in which developing lymphocytes assemble an enormous repertoire of antigen receptor genes (>10^8). These immunoglobulin (Ig) and T-cell receptor (TCR) genes are generated from large arrays of gene segments by a unique process of somatic DNA recombination, called V(D)J recombination. As a result, each precursor B- or T-lymphocyte bears a unique antigen receptor that, following negative selection to delete autoreactive clones, will recognize its signature spectrum of foreign antigens. However, receptor diversification by a process that alters the genome of somatic cells comes at a cost. Genetic defects that compromise any step of the complex V(D)J recombination mechanism block antigen receptor gene assembly and lymphocyte development, resulting in a severe combined immunodeficiency.[1,2] Alternatively, aberrant targeting of the recombination apparatus (recombinase) to regions of the genome harboring oncogenes leads to chromosomal translocations that underlie a majority of lymphoid tumors (e.g., leukemias and lymphomas).[3] Although the basic mechanisms of the V(D)J recombination reaction have been worked out in great detail, we are still unraveling

*Corresponding Author: Eugene M. Oltz—Department of Microbiology and Immunology, Vanderbilt University, Nashville TN, 37232, USA. Email: eugene.oltz@vanderbilt.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

the genetic and epigenetic strategies employed by developing lymphocytes to differentially target recombinase activity at specific regions within antigen receptor loci. In this chapter, we review our current understanding of these regulatory strategies. One emerging theme from recent studies is that only a subset of control mechanisms is broadly employed at all antigen receptor loci.[4] Indeed, a potpourri of regulatory strategies governs the targeting of distinct gene segment clusters by recombinase. Many of the lessons learned through the study of antigen receptor gene control are broadly applicable to the dynamic regulation of complex genetic loci during development, cellular activation and cellular differentiation.

Developmental Control of V(D)J Recombination

The enzymatic components of V(D)J recombinase are products of the Recombination Activating Genes 1 and 2 (RAG-1/2), which are expressed specifically in precursor lymphocytes.[5,6] The RAG-1/2 complex targets recombination signal sequences (RSSs) that flank V, D and J gene segments within all antigen receptor loci.[4] The RSSs are composed of a palindromic heptamer, which abuts each coding segment, a nonconserved spacer of 12 or 23 base pairs (bp) in length and an AT-rich nonamer.[7] The molecular constraints of V(D)J recombination include a strict requirement for synapsis between the RAG-1/2 complex and two RSSs, one of which must harbor a 12 bp spacer and the other a 23 bp spacer (the 12/23 rule). Once the recombinase forms such a synapse, it introduces double-stranded DNA breaks precisely at the junction between the two compatible RSSs and the coding region of their adjacent gene segments. RAG-mediated cleavage thus generates two signal ends, which contain the RSSs and intervening DNA and two coding ends, which contain the gene segments and the remainder of the broken chromosome. The two types of ends are then sealed separately by the cellular DNA repair machinery to create a signal join—usually in the form of an extrachromosomal circle containing the two RSSs—and a coding join, which fuses the two selected gene segments irreversibly in the cellular genome.[8]

Seminal studies from the Alt laboratory in the 1980s showed that a single recombinase, now known to be RAG-1/2, could rearrange both Ig and TCR gene segments in transformed cells.[9] However, in vivo, the process of V(D)J recombination is tightly regulated at three major levels.[4,10] The first and most obvious level is tissue-specificity. Although both precursor B- and T-lymphocytes express RAG-1/2, thymocytes only target TCR genes for recombination while Ig genes are specifically targeted by developing B-cells. Second, all antigen receptor genes are assembled via a stepwise process that is intimately coupled to the lymphocyte developmental program. Upon lineage commitment, pro-B-cells first target recombinase to the DHJH cluster within the Ig heavy chain (IgH) locus (Fig. 1A). Following DH→JH recombination, a pro-B-cell clone will then retarget recombinase to the upstream VH cluster to generate a VHDHJH join.[11] If this coding join is in-frame, the cell expresses IgH protein, which in combination with a surrogate light chain forms a preB-cell receptor (pre-BCR).[12] If the first IgH allele is rearranged out-of-frame, the second allele is then targeted for VH→DHJH recombination. Once formed, the pre-BCR signals for developmental progression of the pro-B-cell to the pre-B stage, which then completes Ig light chain (IgL) gene assembly in an ordered manner (first Vκ→Jκ then Vλ→Jλ if both Igκ alleles are out-of-frame).[13] Thymocytes undergo an analogous developmental program to produce TCRα/β T-cells using the following ordered rearrangement process (Fig. 1B): (i) Dβ→Jβ then (ii) Vβ→DβJβ in $CD4^-CD8^-$ (double-negative, DN) pro-T-cells and (iii) Vα→Jα in $CD4^+CD8^+$ (double-positive, DP) preT-cells.[4,10]

The third level of control is imposed on V(D)J recombination at the developmental transition between pro-B/T and pre-B/T-cells. Following expression of the pre-BCR or pre-TCR, each developing precursor cell must shut down further rearrangement at IgH or TCRβ loci, respectively, despite continued expression of recombinase. This process, called allelic exclusion, precludes the generation of two productive IgH or TCRβ rearrangements in a given cell.[14-16] Allelic exclusion is essential for ensuring the monospecificity of antigen receptors on mature lymphocytes, a cardinal feature of the adaptive immune response. Together, the three levels of control imposed on V(D)J recombination assure expression of the proper antigen receptor by

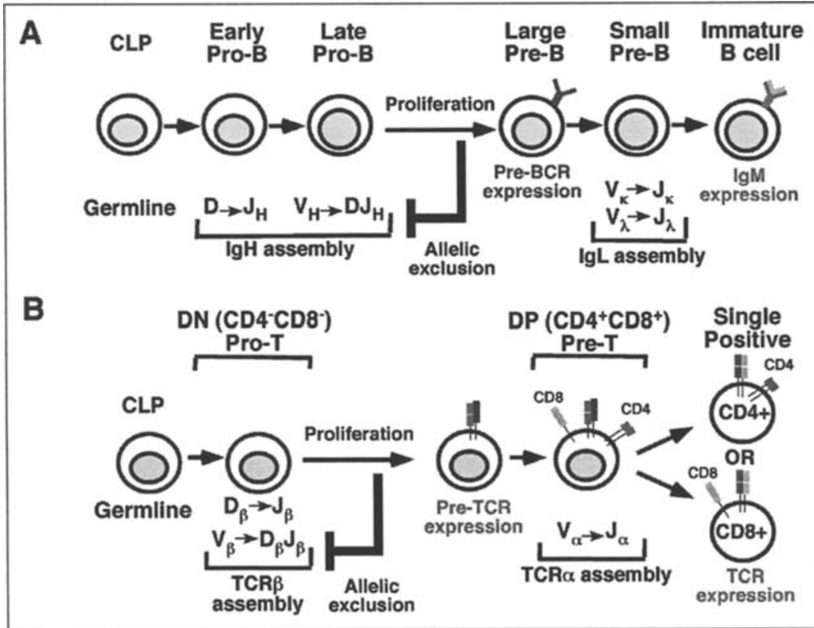


Figure 1. Control of Antigen receptor gene assembly during lymphocyte development. A) Precursor B-cell development and Ig gene assembly. Mouse B-cell development proceeds in the bone marrow initiating from a common lymphoid progenitor (CLP), proceeding through to the pro-B, pre-B and immature B-cell stages. Each developmental step is guided by stage-specific recombination of the B-cell receptor Ig genes and results in the expression of a signature IgM protein on the surface of the immature B-cell. The order in which specific IgH and IgL chain gene rearrangements occur is indicated below each cell type. Pre-BCR expression results in a proliferative burst and feedback inhibition of further IgH gene rearrangement (allelic exclusion). B) Precursor T-cell development and TCR gene assembly. After migration of CLPs from the bone marrow, T-cell development in the thymus proceeds in a stepwise fashion through the pro-T, pre-T and immature T-cell stages. As indicated, each stage is marked by the ordered rearrangement of TCR genes and the surface expression of CD4 and CD8 coreceptor molecules. Pre-TCR receptor expression signifies progression to the preT-cell stage and blocks further TCRβ gene rearrangement.

B versus T-cells, coordinate lymphocyte development programs and maintain a single, unique binding signature for each lymphocyte clone.

Genetic Control of Recombinase Accessibility

Once it was discovered that all precursor lymphocytes employ a single recombinase to target indistinguishable RSSs, a key question became how clusters of Ig and TCR gene segments could be targeted with the observed tissue-, stage- and allele-specificity. An important clue came from the observation that unrearranged Ig gene segments are transcribed at the precise developmental time points they are targeted for rearrangement.[17,18] Since this original observation, the correlation between germline transcription of gene segments and their recombination has been extended to all antigen receptor loci[4,19] and gave birth to the accessibility model for control of V(D)J recombination. In its simplest form, this model invokes changes in the accessibility of gene segments to the recombinase as the key mechanism by which precursor lymphocytes target specific portions of an antigen receptor locus.[20] For example, upon commitment to the T-cell lineage, the DβJβ cluster "opens up" and becomes accessible to RNA polymerase and RAG complexes, whereas the

TCRα and Ig loci remain hidden from these nuclear factors (Fig. 2). The accessibility model has been confirmed by numerous subsequent studies, including those by Schlissel and colleagues who showed that infusion of recombinant RAG complexes into nuclei from precursor lymphocytes cleaves only gene segments that are targeted by endogenous recombinase at a particular developmental stage.[21]

The connection between germline transcription and recombination led Alt and colleagues to test whether common genetic elements control both processes, perhaps by altering locus accessibility.[22,23] At the time, it was known that transcription of all antigen receptor loci is regulated by enhancer elements, usually located at the 3' end of the locus and a series of promoters situated upstream of individual gene segments or 5' to a cluster of related segments (Fig. 2). Targeted deletion of known enhancers from Ig and TCR loci severely inhibits both germline transcription and recombination of gene segments in cis.[4,24-27] These findings support a dual role for enhancers as transcriptional regulatory regions and as accessibility control elements (ACEs) for V(D)J recombination. Subsequent studies demonstrated a similar role for germline promoters, which have a more localized impact on the accessibility of neighboring gene segments (see below).[28-31]

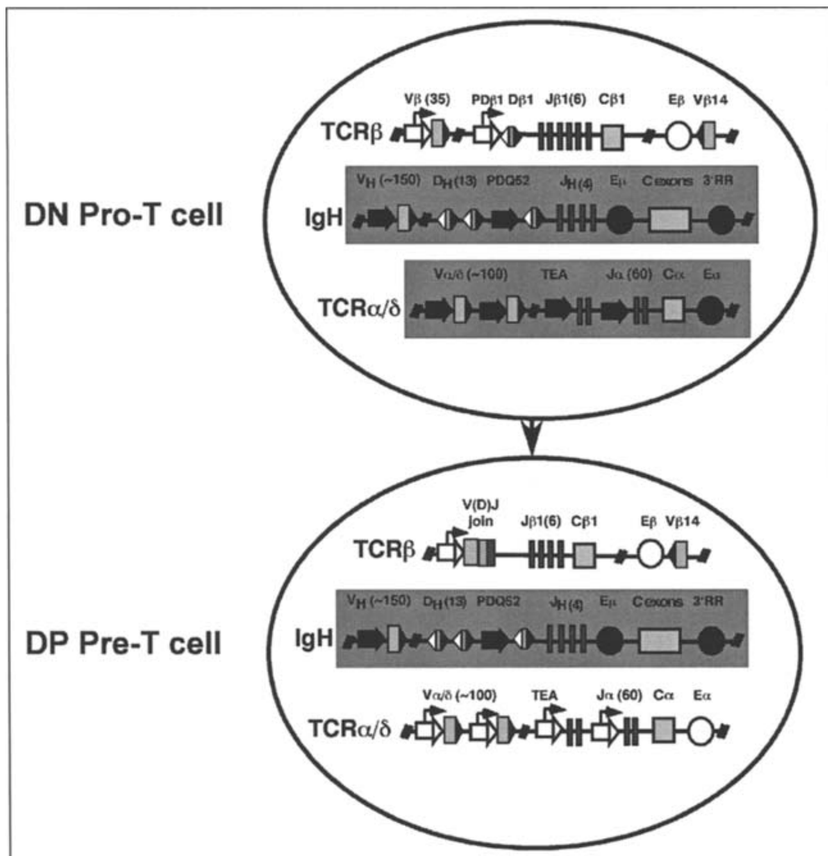


Figure 2. Stage-specific activation of TCR genes. Schematic representation of TCRβ, TCRα/δ and IgH genes are depicted. Block arrows indicate promoter regions and circles represent enhancer elements. The chromatin accessibility status of these regulatory elements at the DN (double-negative) pro-T-ell stage and the DP (double-positive) preT-cell stage is specified (white = "open," gray = "closed"). The shaded boxes overlying the loci indicate a closed configuration.

The precise mechanisms by which ACEs control V(D)J recombination are still under study and may differ significantly between regions within each antigen receptor locus. For example, an obvious mechanism by which promoters might serve as ACEs is via transcriptional readthrough of downstream gene segments, which is known to impact the configuration of associated chromatin. This appears to be the case at the Jα cluster of gene segments, which harbors a series of promoters directing transcription of segments up to several kb away.[32] Disruption of readthrough transcription by placing a strong terminator within a Jα cluster leaves recombination of upstream gene segments intact but severely impairs rearrangement of Jα segments downstream of the transcriptional terminator.[33] In contrast, the Dβ1 germline promoter regulates recombinase accessibility of the small Dβ1Jβ1 cluster (1-2 kb) independent of transcriptional readthrough of the gene segments.[29] Taken together, these studies suggest that transcription may be dispensible for promoter-directed accessibility of proximal gene segments but may significantly augment recombinase accessibility at more distal elements.

Chromatin Accessibility Control Mechanisms for V(D)J Recombination

A major strategy of eukaryotic cells for controlling DNA access to nuclear factors is through changes in the configuration of chromatin. Classical histology studies defined several forms of chromatin that vary in their degree of compaction and, therefore, accessibility to nuclear factors.[34] Heterochromatin is the most compacted form and normally is associated with transcriptionally silent loci. Euchromatin is the most relaxed form and is usually rich in transcribed genes. In between these two extremes lies facultative heterochromatin, which harbors many molecular signatures of silent chromatin, but unlike heterochromatin, is more easily converted to an open state. As such, facultative heterochromatin associates with genes that are under dynamic transcriptional control. The two building blocks of chromatin, nucleosomes and DNA, form a spool-like structure in which each nucleosome is wrapped by the DNA helix.[35,36] Nucleosomes, in turn, are composed of an octamer of histones called H2A, H2B, H3 and H4. Histones H3 and H4 are characterized by an N-terminal tail that is targeted for a broad panel of covalent modifications.

An emerging picture in gene regulation is that covalent modification of histones plays an important role in determining the accessibility status of associated DNA. Indeed, the pattern of histone modifications is now thought to constitute a "code" that is recognized by other nuclear factors to alter local chromatin accessibility.[37] For example, most transcriptionally active genes associate with nucleosomes harboring the following covalent modifications: acetylation of H3-Lysine 9 (H3K9), acetylation of H4K8 and K12 and methylation of H3K4. Acetylated histones attract other nuclear factors containing bromodomains that further augment accessibility, including histone acetyltransferases (HATs) and nucleosome remodeling complexes.[38] Transcriptionally inactive genes associate with chromatin lacking the aforementioned modifications, but instead display H3K9 and H3K27 methylation. These methylation marks are recognized by proteins containing a chromodomain, which facilitate the formation of more highly compacted chromatin (e.g., histone deacetylase—HDACs).[39] In addition to histones, methylation of DNA at CpG dinucleotides correlates with levels of chromatin accessibility and gene activity. In general, CpG dinucleotides are hypermethylated at transcriptionally silent loci but hypomethylated at active genes.[40] The modified CpG dinucleotide is targeted by methyl CpG-binding proteins (e.g., MeCP1), which have been mechanistically linked to gene repression via their association with HDACs and H3K9/K27 methyltransferases.[41]

Given the connection between transcription and V(D)J recombination, it seemed likely that chromatin modifications play an important role in recombinase accessibility. In pioneering studies, Krangel and colleagues proved this hypothesis to be correct. They showed that H3K9 acetylation associates with TCRα/δ gene segments when they are actively undergoing recombination.[42] Since this original report, numerous groups have shown that an identical pattern of histone and CpG modifications characterize loci that are transcriptionally or recombinationally active and a separate set of modifications decorate inert antigen receptor loci (Fig. 3).[4,43] Importantly, ACE deletions

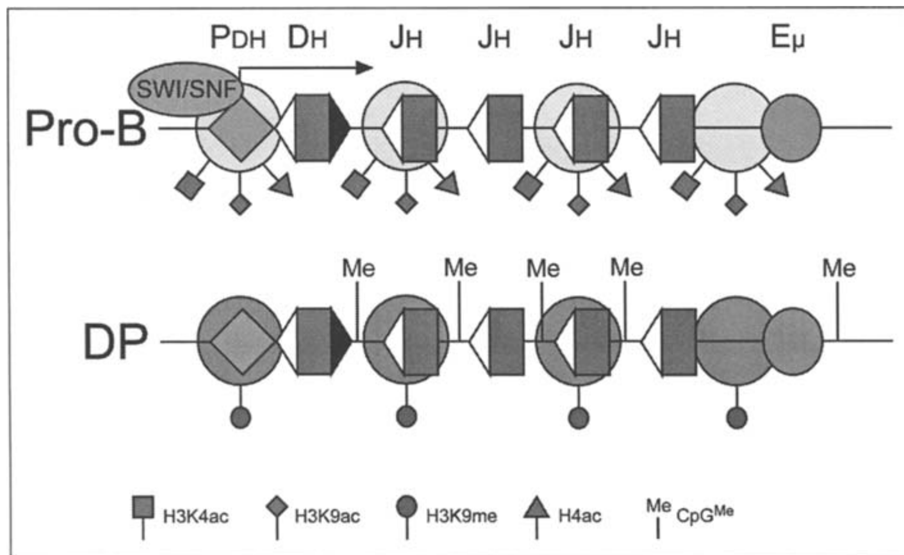


Figure 3. Epigenetic regulation of recombinase accessibility. Schematic representation of histone modifications at the mouse heavy chain locus in pro-B and DP pro-T-cells. RSSs are represented as triangles (23 bp spacer black and 12 bp spacer white) flanking each DH and JH gene segment (rectangles). Chromatin associated with the DHJH cluster in pro-B-cells is marked with an "accessibile" pattern of modifications (H3K4me, H3K9ac and H4ac). The germline promoter (diamond) together with acetylated histones recruit the SWI/SNF complex (oval), which remodels chromatin at neighboring gene segments to generate a recombinase accessible state in pro-B-cells. In DP thymocytes, the DHJH cluster is decorated with H3K9me and CpG methylation, two chromatin modifications that mediate repression.

from Ig or TCR loci convert histone modifications from an "accessible" to an "inaccessible" pattern.[28,44] In recent studies, we established a cause-effect relationship for one histone modification in the control of V(D)J recombination.[45] Specifically, we showed that recruitment of an H3K9 methyltransferase inhibited recombination of stably integrated substrates, despite the presence of requisite ACEs. Thus, H3K9 methylation is sufficient in this setting to repress recombinase accessibility. In n important new development, Oettinger and colleagues have shown that the plant homeodomain of RAG-2 binds specifically to tri-methylated H3K4, a characteristic mark of transcriptionally active chromatin.46,47 Mutations in RAG-2 that disrupt H3K4me3 binding also attenuate recombinase activity, indicating that the interaction is functionally significant. Thus, the pattern of histone modifications at a given cluster of gene segments may fine-tune its affinity for recombinase, providing a level of control beyond sheer accessibility. Despite these advances, a cause-effect relationship has yet to be established for any histone modification and the recombination potential of gene segments at an endogenous antigen receptor locus.

An added layer of complexity to the "histone code" derives from the large collection of nuclear factors that can potentially recognize each covalent modification. In this regard, an important role for H3 and H4 acetylation is thought to be the docking of nucleosome remodeling complexes via bromodomains.[38] In vitro, these complexes alter the conformation or position of nucleosomes on DNA, thereby altering the accessibility of neighboring sites to nuclear factors.[49] Certain complexes, such as SWI/SNF, augment accessibility, whereas other remodeling machines normally induce a higher degree of compaction.[50] At the heart of each remodeling complex is an ATP-dependent motor that provides the energy for reconfiguring nucleosome conformation or sliding the nucleosome to a new position. In the case of SWI/SNF, the major ATP-dependent

subunits are called Brg1 and Brm. In the context of antigen receptor gene assembly, early studies on chromatin remodeling were largely restricted to in vitro substrates. In one such trailblazing study, Oettinger and colleagues showed that Brg1 could reverse the block to recombinase accessibility when added to nucleosomal substrates.[51] This group went on to show that Brg1 associates with Ig and TCR loci only when they are poised to undergo rearrangement and are therefore in a recombinase-accessible state.[43] Most recently, we have triangulated the relationship between ACE function, SWI/SNF association and recombinase accessibility. Recruitment of Brg1 to the endogenous TCRβ locus requires both the Dβ1 germline promoter and Eβ enhancer elements.[52] In chromosomal substrates, ACE function of the germline promoter can be replaced completely by artificial recruitment of Brg1. Most importantly, RNAi-mediated ablation of Brg1 and Brm in primary thymocytes abrogates both germline transcription and Dβ→Jβ recombination. Together with prior studies, these findings indicate that ACEs likely function to alter histone modification patterns within antigen receptor loci in order to recruit chromatin remodeling complexes that either impart or impair recombinase accessibility (Fig. 3).

Control of V(D)J Recombination by Nuclear Compartmentalization

Regulation of gene accessibility is a complex process that involves not only the covalent modification of histones and DNA, but also the localization of large genetic loci to distinct regions within the cell nucleus. In general, transcriptionally silent genes are located near the nuclear periphery, whereas expressed genes reside at a more central location within the nucleus.[53] Although the underlying mechanisms for this effect remain to be established, recent experiments demonstrate that enforced compartmentalization of a genetic locus to the inner nuclear membrane represses its expression.53 With regards to antigen receptor loci, fluorescence in situ hybridization (FISH) analysis has provided insights into how subnuclear relocation may influence their assembly by V(D)J recombination during lymphocyte development. In pro-B-cells poised to initiate DH→JH recombination, both the IgH and Igκ alleles are preferentially positioned in the central portion of the nucleus. In contrast, these loci occupy a perinuclear location in hematopoietic progenitors and pro-T-cells, which do not rearrange Ig genes.[54] Thus, tissue-specific activation of Ig loci may involve their repositioning from the nuclear periphery in hematopoietic stem cells to a more central location upon commitment to the pro-B-cell stage (see below and Fig. 5).

Primary Activation of Antigen Receptor Loci for D to J Rearrangement

V(D)J recombination at IgH and TCRβ loci progresses in a step-wise fashion, initiating with the assembly of a DJ join, followed by rearrangement of a V gene segment to the existing DJ element. The genetic and epigenetic mechanisms leading to the crucial first step, activating DJ clusters for rearrangement, have been extensively studied for the TCRβ locus. Due to its compartmentalized architecture, this antigen receptor locus serves as a tractable model to study the precise role of ACEs in tissue-specific activation of D→J recombination. The mouse TCRβ locus consists of two distinct DβJβ clusters, which are both under the control of a single 3' enhancer element (Eβ). Eβ function is specific for T lineage cells and this control element becomes activated at the earliest stage of thymocyte development.[55] In addition, both DβJβ clusters harbor a single germline promoter located proximal to each Dβ gene segment (PDβ1 and PDβ2), which are completely dependent on Eβ for their function (Fig. 4).[56,57] Knockout studies in mice have shown that Eβ provides a long-range ACE function to direct chromatin modifications, recombination and germline transcription at both DβJβ clusters.[24,25,44] In contrast, targeted deletion of the germline promoter located directly upstream of the Dβ1 gene segment cripples rearrangement and transcription of the Dβ1Jβ1 but not the Dβ2Jβ2 cluster, indicating that germline promoters have a more local influence on chromatin accessibility.[31]

Although the cis- and trans-elements involved in recombination are unique for individual loci, studies of TCRβ have provided a mechanistic model for initial activation of DJ clusters for their targeting by V(D)J recombinase (Fig. 4).[28] First, thymocyte precursors activate an inherent ACE

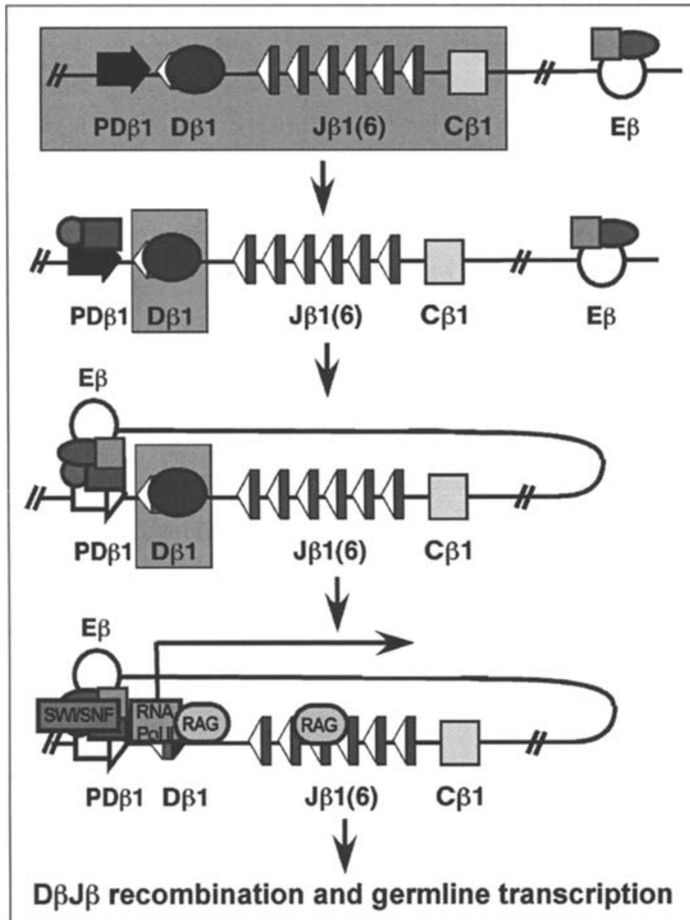


Figure 4. Ordered activation of Dβ→Jβ rearrangement. Initially, thymocyte precursors activate the T-cell specific enhancer, Eβ (circle) via the binding of transcription factors (shaded shapes). The ACE function of Eβ mediates a spread of chromatin accessibility throughout the majority of the DβJβ cluster, with the exception of the Dβ1-proximal region, which remains refractory to this opening (gray shaded box and associated nucleosome, gray oval). The germline promoter, PDβ1, becomes activated and binds transcription factors in an Eβ-dependent manner. Once activated, these two distal regulatory elements interact to form a PDβ/Eβ holocomplex, which in turn recruits chromatin remodeling complexes, including SWI/SNF. SWI/SNF remodels nucleosomes associated with the Dβ1 region, exposing the RSS and TATA box for Dβ→Jβ recombination and germline transcription.

function of the Eβ enhancer, which mediates the spread of partially accessible chromatin over a large region spanning both DβJβ clusters. Although the mechanisms are currently unknown, chromatin associated with the Dβ1 gene segment is initially refractory to Eβ-mediated opening.[28] However, enhancer-mediated reorganization of TCRβ chromatin permits binding of additional transcription factors to the Dβ1 germline promoter.[44] The loaded promoter and distal Eβ element physically interact, presumably via factors bound to each ACE, generating a stable holocomplex.[28] The promoter/enhancer holocomplex may serve as a staging platform to recruit SWI/SNF, which in turn reorganizes local nucleosomes, especially those associated with the Dβ1-RSS.[52] The SWI/

SNF-mediated remodeling likely exposes the underlying TATA box and RSS to initiate germline transcription and Dβ→Jβ recombination, respectively. Future studies should be directed at understanding the generality of this contingent mechanism for recombinase accessibility at the IgH and other antigen receptor loci.

Long-Range Control of V(D)J Recombination

Changes in chromatin accessibility play a key role in controlling recombination of D and J segments, which are separated by relatively short distances in the TCRβ and IgH loci (<20 kb). In contrast, the V clusters are separated by much longer stretches of DNA from their partner D and/or J segments (>100kb) at the IgH, TCRβ, Igκ, Igλ and TCRα loci. As such, additional mechanisms may be required to facilitate efficient V→(D)J recombination. Whereas many of the control elements that regulate D→J recombination have been identified, the ACEs guiding V→DJ remain largely unknown. Knockout experiments that delete known enhancer and germline D promoter elements have no effect on the transcription or accessibility of distal V gene segments at either the IgH or TCRβ locus.[57-59] Emerging evidence points to a requirement for Vβ or VH promoters in modulating local chromatin accessibility to recombinase. Deletion of the Vβ13 promoter cripples germline transcription and rearrangement of that gene segment.[60] However, reporter gene assays indicate that transcription from Vβ and VH promoters is enhancer-dependent. Thus, the epigenetic regulation of the V gene clusters is under the control of unidentified enhancers that reconfigure chromatin to mediate their recombinase accessibility.

Despite a clear requirement for changes in chromatin accessibility to trigger long-range V→(D)J recombination, epigenetic changes alone are not sufficient. Introduction of the TCRα enhancer proximal to a Vβ gene segment greatly augments its chromatin accessibility and germline transcription in DP thymocytes, a developmental stage where Vβ segments are repressed by allelic exclusion. However, enhanced chromatin accessibility is insufficient to drive Vβ→DβJβ recombination involving the targeted gene segment.[61] Thus, additional mechanisms are required to mediate the long-range interactions between distal V arrays and downstream (D)J regions.

Insights into the mechanisms that control chromosomal dynamics within the TCRβ locus derive from studies using three-dimensional FISH. These analyses produced astounding links between changes in subnuclear location, topography and TCRβ gene regulation (Fig. 5).[16] Through a mechanism called locus contraction, the distal Vβ cluster loops to become spatially proximal to the DβJβ cluster in DN thymocytes but not in other tissues.[62] The changes in TCRβ locus topology were confirmed using molecular analysis, which revealed physical associations not only between Vβ and DβJβ clusters but also amongst Vβ gene segments themselves. A similar contraction process regulates long-range interactions of IgH and Igκ loci, which span distances of 2.5 megabases and 3 megabases, respectively. Three-dimensional FISH revealed that the VH region of IgH is juxtaposed with the DHJH domain via a looping mechanism, which occurs specifically at the pro-B-cell stage of development.[54,63,64] Likewise, the Igκ locus undergoes active contraction in preB-cells to bring Vκ and Jκ gene segments into spatial proximity.[65] Thus, changes in locus topology appear to be a general mechanism for long-range control of recombination between distant V and (D)J gene segments. Together, these findings suggest that novel cis-elements within both the V and the (D) J clusters mediate their physical association, perhaps forming an active hub where multiple V gene segments are in spatial proximity to their partner (D)J segments (Fig. 5).

Allelic Exclusion

The specificity of immune responses is maintained by restricting each lymphocyte clone to express a single antigen receptor gene combination. Following a productive rearrangement on one Ig or TCR allele, precursor lymphocytes must rapidly inhibit rearrangement at the second allele, a process called allelic exclusion. Indeed, transgenic mice engineered to express functional IgH or TCRβ proteins in precursor lymphocytes freeze V→DJ rearrangement at their corresponding endogenous alleles.[66,67] These findings indicate that the pre-BCR and pre-TCR complexes participate in a feedback mechanism that specifically disrupts long-range V→(D)J recombination.

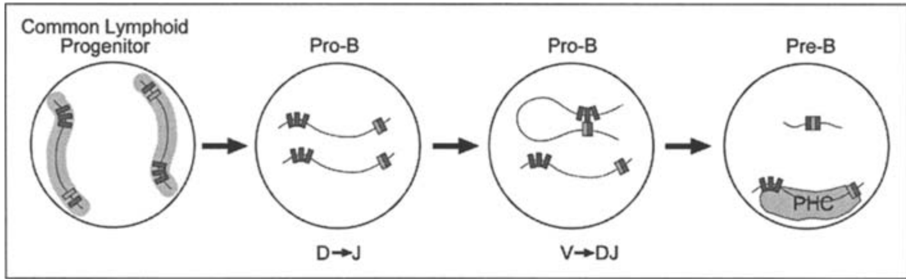


Figure 5. Regulation of V(D)J recombination by chromosomal dynamics. Cartoon showing changes in nuclear positioning and IgH locus topology during B-cell development. Prior to B lineage commitment, both IgH alleles are positioned at the nuclear periphery and remain inaccessible to recombinase in common lymphoid progenitors. Upon lineage commitment, both IgH loci migrate to a more central location in the pro-B-cell nucleus and undergo DH→JH recombination. IgH loci also contract via a looping mechanism to position the VH cluster into spatial proximity with the DHJH fusions, permitting efficient VH→DHJH recombination. If the first IgH allele undergoes a productive rearrangement, preBCR signaling mediates a repositioning of the second, DHJH rearranged allele to associate with pericentromeric heterochromatin (PHC).

Although the mechanisms are poorly understood, allelic exclusion is controlled at multiple levels including (i) chromatin accessibility, (ii) locus topology and (iii) repositioning of loci to pericentromeric heterochromatin. With regards to chromatin accessibility, hyperacetylation of histones associated with Vβ gene segments is lost following TCRβ expression and the transition of thymocytes from the DN to DP stage of development.[68] Similar changes in the pattern of histone modifications suggest a loss of chromatin accessibility at VH gene segments upon expression of IgH proteins.[69] However, as stated above, allelic exclusion is still enforced at a Vβ gene segment despite the neighboring insertion of a functional ACE that opens associated chromatin in DP thymocytes.[61]

Similar to activation of V→(D)J rearrangement, changes in locus topology may play an important role in shutting down this stage of antigen receptor gene assembly to enforce allelic exclusion. FISH analyses of thymocyte populations revealed that juxtaposition of the Vβ and DβJβ regions persists throughout the DN stage but the locus "decontracts" and becomes linear at the DP stage, where Vβ→DβJβ recombination is suppressed by allelic exclusion.[62] A similar decontraction process takes place at the IgH locus during the pro-B to preB-cell transition.[63] In addition to distancing the V and (D)J clusters by decontraction, the nonfunctional IgH and TCRβ alleles reposition themselves to associate with pericentromeric heterochromatin, a potently repressive environment (Fig. 5). Recent studies indicate that IgH decontraction and its association with heterochromatin are mediated by the Igκ 3' enhancer, which in some manner directs pairing between the IgH and Igκ loci in pre-B cells.[72] Thus, in addition to chromatin accessibility, allelic exclusion of V→(D)J recombination is likely orchestrated by wholescale changes in the structure of the nonfunctional locus, which (i) separate and thereby isolate the V and (D)J clusters and (ii) move the gene segments into a repressive chromatin environment that is refractory to nuclear factors such as recombinase.

Allelic exclusion at the IgL loci incorporates many of the aforementioned regulatory strategies but also involves two additional levels of control. Both Igκ alleles migrate from the nuclear periphery to a central location in pro-B-cells.[65] However, upon developmental transition to the preB-cell stage a single Igκ allele becomes activated based on chromatin modifications, its early replication, germline Jκ transcription and Vκ→Jκ recombination.[70] The second Igκ allele initially associates with pericentromeric heterochromatin, which represses its recombinase accessibility. Presumably the second allele dissociates from this repressive environment over time if recombination of the first

allele fails to generate a functional BCR.[65] Thus, unlike IgH or TCRβ loci, allelic exclusion at Igκ involves sequential activation of the two separate alleles. Finally, once a functional, alloreactive BCR is produced by a pre-B-cell, V(D)J recombination is terminated via a signal-dependent repression of RAG expression, thus providing an irreversible means to enforce allelic exclusion.[71]

Conclusion

During the past five years we have witnessed an explosion in our understanding of mechanisms that control the targeting of recombinase to precise regions within antigen receptor loci and thereby provide requisite tissue-, stage- and allele-specificity. We are beginning to appreciate that, although common strategies are employed at many loci, the activation or repression of recombination at each cluster of gene segments likely proceeds via distinct mechanisms. Indeed, each region appears to use a unique combination of crosstalk between ACE function, epigenetic modifications, transcription, chromatin remodeling, nuclear localization and locus contraction to achieve the required level of recombination efficiency. Additional mechanisms will likely be discovered in the near future. We have begun to move forward from correlative studies focused on the molecular hallmarks of recombinase accessibility to a clearer definition of whether these features are causally linked. However, similar advances must be made to establish cause-effect relationships between long-range changes in locus conformation (e.g., contraction) and recombination. Of equal importance, a considerable effort should be placed in the discovery of cis-acting elements that control long-range interactions between distant clusters of gene segments. Given what we have already learned from antigen receptor gene regulation, we can anticipate that many of these new findings will have broad implications for the control of gene expression.

Acknowledgements

The authors acknowledge support received from the National Institutes of Health.

References

1. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell 2002; 109 Suppl:S45-55.
2. de Villartay JP, Poinsignon C, de Chasseval R et al. Human and animal models of V(D)J recombination deficiency. Curr Opin Immunol 2003; 15(5):592-598.
3. Jankovic M, Nussenzweig A, Nussenzweig MC. Antigen receptor diversification and chromosome translocations. Nat Immunol 2007; 8(8):801-808.
4. Cobb RM, Oestreich KJ, Osipovich OA et al. Accessibility control of V(D)J recombination. Adv Immunol 2006 91:45-109.
5. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248(4962):1517-1523.
6. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell 1989; 59(6):1035-1048.
7. Sakano H, Huppi K, Heinrich G et al. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 1979; 280(5720):288-294.
8. Dudley DD, Chaudhuri J, Bassing CH et al. Mechanism and Control of V(D)J Recombination versus Class Switch Recombination: Similarities and Differences. Adv Immunol 2005; 86:43-112.
9. Yancopoulos GD, Blackwell TK, Suh H et al. Introduced T-cell receptor variable region gene segments recombine in preB cells: evidence that B- and T-cells use a common recombinase. Cell 1986; 44(2):251-259.
10. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
11. Hardy RR, Carmack CE, Shinton SA et al. Resolution and characterization of pro-B and prepro-B-cell stages in normal mouse bone marrow. J Exp Med 1991; 173(5):1213-1225.
12. Melchers F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. Nat Rev Immunol 2005; 5(7):578-584.
13. Ehlich A, Schaal S, Gu H et al. Immunoglobulin heavy and light chain genes rearrange independently at early stages of B-cell development. Cell 1993; 72(5):695-704.
14. Krangel MS. T-cell development: better living through chromatin. Nat Immunol 2007; 8(7):687-694.
15. Loffert D, Ehlich A, Muller W et al. Surrogate light chain expression is required to establish immunoglobulin heavy chain allelic exclusion during early B-cell development. Immunity 1996; 4(2):133-144.

16. Sen R, Oltz E. Genetic and epigenetic regulation of IgH gene assembly. Curr Opin Immunol 2006; 18(3):237-242.
17. Van Ness BG, Weigert M, Coleclough C et al. Transcription of the unrearranged mouse Cκ locus: sequence of the initiation region and comparison of activity with a rearranged Vκ- Cκ gene. Cell 1981; 27(3):593-602.
18. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. Cell 1985; 40(2):271-281.
19. Sleckman BP, Gorman JR, Alt FW. Accessibility control of antigen-receptor variable-region gene assembly: role of cis-acting elements. Annu Rev Immunol 1996; 14:459-481.
20. Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. Annu Rev Immunol 1986; 4:339-368.
21. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85(6):887-897.
22. Ferrier P, Krippl B, Blackwell TK et al. Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. EMBO J 1990; 9(1):117-125.
23. Oltz EM, Alt FW, Lin WC et al. A V(D)J recombinase-inducible B-cell line: role of transcriptional enhancer elements in directing V(D)J recombination. Mol Cell Biol 1993; 13(10):6223-6230.
24. Bories JC, Demengeot J, Davidson L et al. Gene-targeted deletion and replacement mutations of the T-cell receptor beta-chain enhancer: the role of enhancer elements in controlling V(D)J recombination accessibility. Proc Natl Acad Sci USA 1996; 93(15):7871-7876.
25. Bouvier G, Watrin F, Naspetti M et al. Deletion of the mouse T-cell receptor beta gene enhancer blocks αβ T-cell development. Proc Natl Acad Sci USA 1996; 93(15):7877-7881.
26. Chen J, Young F, Bottaro A et al. Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. EMBO J. 1993; 12(12):4635-4645.
27. Serwe M, Sablitzky F. V(D)J recombination in B-cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. EMBO J. 1993; 12(6):2321-2327.
28. Oestreich KJ, Cobb RM, Pierce S et al. Regulation of TCRβ gene assembly by a promoter/enhancer holocomplex. Immunity 2006; 24(4):381-391.
29. Sikes ML, Meade A, Tripathi R et al. Regulation of V(D)J recombination: A dominant role for promoter positioning in gene segment accessibility. Proc Natl Acad Sci USA 2002; 99:12309-12314.
30. Sikes ML, Suarez CC, Oltz EM. Regulation of V(D)J recombination by transcriptional promoters. Mol Cell Biol 1999; 19(4):2773-2781.
31. Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the Dβ1 gene segment at the TCRβ locus by a germline promoter. Immunity 1999; 10(3):313-322.
32. Hawwari A, Bock C, Krangel MS. Regulation of T-cell receptor alpha gene assembly by a complex hierarchy of germline Jα promoters. Nat Immunol 2005; 6(5):481-489.
33. Abarrategui I, Krangel MS. Regulation of T-cell receptor-alpha gene recombination by transcription. Nat Immunol 2006; 7(10):1109-1115.
34. Fahrner JA, Baylin SB. Heterochromatin: stable and unstable invasions at home and abroad. Genes Dev 2003; 17(15):1805-1812.
35. Khorasanizadeh S. The nucleosome: from genomic organization to genomic regulation. Cell 2004; 116(2):259-272.
36. Wolffe AP, Guschin D. Review: chromatin structural features and targets that regulate transcription. J Struct Biol 2000; 129(2-3):102-122.
37. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293(5532):1074-1080.
38. Berger SL. Histone modifications in transcriptional regulation. Curr Opin Genet Dev 2002; 12(2):142-148.
39. Lachner M, O'Carroll D, Rea S at al. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 2001; 410(6824):116-120.
40. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002; 16(1):6-21.
41. Fuks F, Hurd PJ, Deplus R et al. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res 2003; 31(9):2305-2312.
42. McMurry MT, Krangel MS. A role for histone acetylation in the developmental regulation of VDJ recombination. Science 2000; 287(5452):495-498.
43. Morshead KB, Ciccone DN, Taverna SD et al. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci USA 2003; 100(20):11577-11582.
44. Spicuglia S, Kumar S, Yeh JH et al. Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. Mol Cell 2002; 10:1479-1487.
45. Osipovich O, Milley R, Meade A et al. Targeted inhibition of V(D)J recombination by a histone methyltransferase. Nat Immunol 2004; 5(3):309-316.

46. Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A PHD-Finger Domain in RAG-2 That Binds Hypermethylated Lysine 4 of Histone H3 Is Necessary for Efficient V(D)J Rearrangement. Immunity 2007; 27(4):561-571.
47. Matthews AG, Kuo AJ, Ramon-Maiques S, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450(7172):1106-1110.
48. Shi X, Hong T, Walter KL et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature 2006; 442(7098):96-99.
49. Kingston RE, Narlikar GJ. ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev 1999; 13(18):2339-2352.
50. Emerson BM. Specificity of gene regulation. Cell 2002; 109(3):267-270.
51. Kwon J, Morshead KB, Guyon JR et al. Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA. Mol Cell 2000; 6(5):1037-1048.
52. Osipovich O, Milley Cobb R, Oestreich KJ et al. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrβ genes. Nat Immunol 2007; 8(8):809-816.
53. Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 2008; 452(7184):243-247.
54. Kosak ST, Skok JA, Medina KL et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 2002; 296(5565):158-162.
55. McDougall S, Peterson CL, Calame K. A transcriptional enhancer 3' of Cβ2 in the T-cell receptor beta locus. Science 1988; 241(4862):205-208.
56. McMillan RE, Sikes ML. Differential Activation of Dual Promoters Alters Dβ2 Germline Transcription during Thymocyte Development. J Immunol 2008; 180(5):3218-3228.
57. Afshar R, Pierce S, Bolland DJ et al. Regulation of IgH gene assembly: role of the intronic enhancer and 5'DQ52 region in targeting DHJH recombination. J Immunol 2006; 176(4):2439-2447.
58. Mathieu N, Hempel WM, Spicuglia S et al. Chromatin remodeling by the T-cell receptor (TCR)-beta gene enhancer during early T-cell development: Implications for the control of TCR-beta locus recombination. J Exp Med 2000; 192(5):625-636.
59. Perlot T, Alt FW, Bassing CH et al. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci USA 2005; 102(40):14362-14367.
60. Ryu CJ, Haines BB, Lee HR et al. The T-cell receptor beta variable gene promoter is required for efficient Vβ rearrangement but not allelic exclusion. Mol Cell Biol 2004; 24(16):7015-7023.
61. Jackson A, Kondilis HD, Khor B et al. Regulation of T-cell receptor beta allelic exclusion at a level beyond accessibility. Nat Immunol 2005; 6(2):189-197.
62. Skok JA, Gisler R, Novatchkova M et al. Reversible contraction by looping of the Tcrα and Tcrβ loci in rearranging thymocytes. Nat Immunol 2007; 8(4):378-387.
63. Roldan E, Fuxa M, Chong W et al. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol 2005; 6(1):31-41.
64. Jhunjhunwala S, van Zelm MC, Peak MM, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell 2008; 133(2):265-279.
65. Goldmit M, Ji Y, Skok J et al. Epigenetic ontogeny of the Igκ locus during B-cell development. Nat Immunol 2005; 6(2):198-203.
66. Gorman JR, Alt FW. Regulation of immunoglobulin light chain isotype expression. Adv Immunol 1998; 69:113-181.
67. Manz J, Denis K, Witte O et al. Feedback inhibition of immunoglobulin gene rearrangement by membrane mu, but not by secreted mu heavy chains. J Exp Med 1988; 168(4):1363-1381.
68. Tripathi R, Jackson A, Krangel MS. A change in the structure of Vβ chromatin associated with TCRβ allelic exclusion. J Immunol 2002; 168(5):2316-2324.
69. Chowdhury D, Sen R. Transient IL-7/IL-7R Signaling Provides a Mechanism for Feedback Inhibition of Immunoglobulin Heavy Chain Gene Rearrangements. Immunity 2003; 18:229-241.
70. Mostoslavsky R, Singh N, Tenzen T et al. Asynchronous replication and allelic exclusion in the immune system. Nature 2001; 414(6860):221-225.
71. Hardy RR, Hayakawa K. B-ell development pathways. Annu Rev Immunol 2001; 19:595-621.

Chapter 10

Molecular Genetics at the T-Cell Receptor β Locus:

Insights into the Regulation of V(D)J Recombination

Marie Bonnet, Pierre Ferrier and Salvatore Spicuglia*

Abstract

The V(D)J recombination machinery assembles antigen receptor genes from germline V, D and J segments during lymphocyte development. In αβT cells, this leads to the production of the T-cell receptor (TCR) α and β chains. Notably, V(D)J recombination at the *Tcrb* locus is tightly controlled at various levels, including cell-type and stage specificities, intralocus ordering and allelic exclusion. Although many of these controls are partly mediated at the level of genomic accessibility to the V(D)J recombinase, recent studies have uncovered novel mechanisms that are also likely to contribute to the developmental regulation of *Tcrb* gene rearrangement events. In this chapter, we summarize our current knowledge and highlight unanswered questions regarding the regulation of V(D)J recombination at the *Tcrb* locus, placing emphasis on mouse transgenesis and gene-targeting approaches.

Introduction

B and T-lymphocytes, the adaptive arms of the immune system in vertebrates, can generate specific responses to a tremendous number of foreign antigens.[1] This remarkable property largely depends on a series of DNA rearrangements between germline V, D and J gene segments that generate variable region genes at antigen receptor-encoding loci. V(D)J recombination events are initiated in the developing lymphocytes by the lymphoid-specific proteins, RAG1 and RAG2 (RAG1/2), which introduce DNA double-strand breaks precisely at the borders between two coding segments and their flanking recombination signal sequences (RSSs). The RSSs are composed of relatively well conserved heptamer and nonamer sequences separated by a less well conserved spacer of either 12 or 23 base-pairs (bp), named 12- and 23-RSSs respectively. V(D)J recombination occurs primarily between one gene segment flanked by a 12-RSS and another by a 23-RSS, a restriction termed the 12/23 rule. Eventually, the broken ends are repaired by the ubiquitously expressed 'nonhomologous end joining' (NHEJ) machinery to form coding and RSS (signal) joints.

There are seven different antigen receptor loci.[1] These include the immunoglobulin heavy (*Igh*) and light (*Igk* and *Igl*) chain loci for B-cells and the T-cell receptor (*Tcr*) *a, b, g* and *d* loci for T-cells. *Tcrb* and *a* gene assembly is carried out at two distinct stages of αβT-cell development in the thymus.[2] V(D)J recombination at the *Tcrb* locus is initiated within the $CD4^-CD8^-$ double-negative (DN) compartment and proceeds stepwise with Dβ-to-Jβ occurring first in DN2 cells, prior to Vβ-to-DJβ joining mostly occurring in DN3 cells. Expression of a functionally rearranged *Tcrb* gene leads thymocytes to pass through β-selection and differentiate into $CD4^+CD8^+$ double-positive

*Corresponding Author: Salvatore Spicuglia—Centre d'Immunologie de Marseille-Luminy, Case 906, Campus de Luminy, Marseille, France. Email: spicuglia@ciml.univ-mrs.fr

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

(DP) cells, while instigating *Tcra* gene rearrangements. Ultimately, TCRαβ-expressing cells may be selected into mature CD4$^+$ or CD8$^+$ single-positive (SP) T-cells.

T-cell development requires temporally-regulated rearrangement and expression of the *Tcr* genes. The lymphoid-specific expression of the *Rag1/2* genes restricts V(D)J recombination to the developing lymphocytes.[1] However, *Rag1/2* regulated expression alone cannot explain all the controls of V(D)J recombination at the *Ig* and *Tcr* endogenous loci. In particular, the *Tcrb* locus is subjected to many levels of regulation which determine a precise developmental order of rearrangement events and ensure that the *Tcrb* gene is expressed in an allelically excluded manner.[2] The present chapter focuses on our current understanding of how V(D)J recombination is regulated at the *Tcrb* locus and places emphasis on mouse transgenesis and gene-targeting approaches that have revealed essential roles for *cis-* and *trans-*regulators of *Tcrb* gene rearrangements and allelic exclusion.

Overview of the *Tcrb* Genomic Structure and Recombination Properties

In the mouse germline, the approximately 700-kilobases (kb) *Tcrb* locus consists of a large (~425-kb) 5' region containing 22 functional Vβ gene segments, as well as 13 additional Vβ pseudogenes and a shorter (~25-kb) 3' region comprising a duplicated cluster of Dβ-Jβ-Cβ gene segments (Fig. 1A).[3] In addition, it contains two groups of trypsinogen genes (not transcribed in T-cells), including one spread over the 250-kb region separating the Vβ and Dβ/Jβ regions. A distinct Vβ gene segment (Vβ14) is situated at the 3' end of the locus, lying in the opposite transcriptional direction. As determined by their RSS types and orientations, recombination of all 5'Vβ, Dβ and Jβ gene segments is deletional, whereas that of Vβ14 occurs by inversion.[3] Germline Vβ and Dβ segments are flanked by upstream transcriptional promoters that initiate sterile transcription in a developmentally regulated fashion. A single transcriptional enhancer (Eβ) lies between Cβ2 and Vβ14. The formation of a complete VDJβ variable region places the promoter of the rearranged Vβ segment in the vicinity and under the control of Eβ, enhancing transcription of the newly-assembled TCRβ unit. Via standard mRNA-maturation processes, the variable VDJβ and constant Cβ exons are then spliced to produce the full-length TCRβ chain. As described further below, the particular structure and organization of the *Tcrb* locus impose a number of constraints regarding the regulation of V(D)J recombination. In this context, it is worth noting that defects in *Tcrb* locus recombination are suspected to contribute to T-cell pathogenesis (e.g., oncogenesis and, not as yet disproven, autoimmunity).[4-6]

Tcrb-RSSs and Rearrangement Efficiency

Despite their overall conservation, *Tcrb*-RSSs exhibit marked sequence variations compared to the canonical RSSs.[3,7,8] Indeed, the frequency of Jβ gene segment usage at the murine *Tcrb* locus correlates well with sequence variations within the corresponding Jβ-12RSS residues.[9] In addition, mutagenesis studies have identified various residues in the sequences of the *Tcrb*-RSSs and their coding flanks that cooperate to determine the ultimate efficiency of the recombination process.[10] Thus, the TCRβ repertoire naturally reflects the subtle interplay between the RSSs and the flanking coding sequences to direct the activity of the V(D)J recombinase. A dramatic, recently uncovered example of this paradigm is illustrated below.

Beyond the 12/23 Rule

As mentioned above, *Tcrb* variable regions are assembled via an ordered two-step process in which Dβ-to-Jβ rearrangement occurs before the appendage of a Vβ to the rearranged DJβ gene segment. In this context, the organization of this locus, in which Vβ and 3'Dβ RSSs contain 23-bp spacers and 5'Dβ and Jβ RSSs contain 12-bp spacers (Fig. 1B), seemed to represent a potential problem since direct joining of Vβ segments to nonrearranged Dβ or Jβ gene segments would also be compatible with the 12/23 rule. However, such rearrangements are respectively very infrequent and practically nonexistent within the endogenous *Tcrb* locus (reviewed by Tillman et al[8]). These

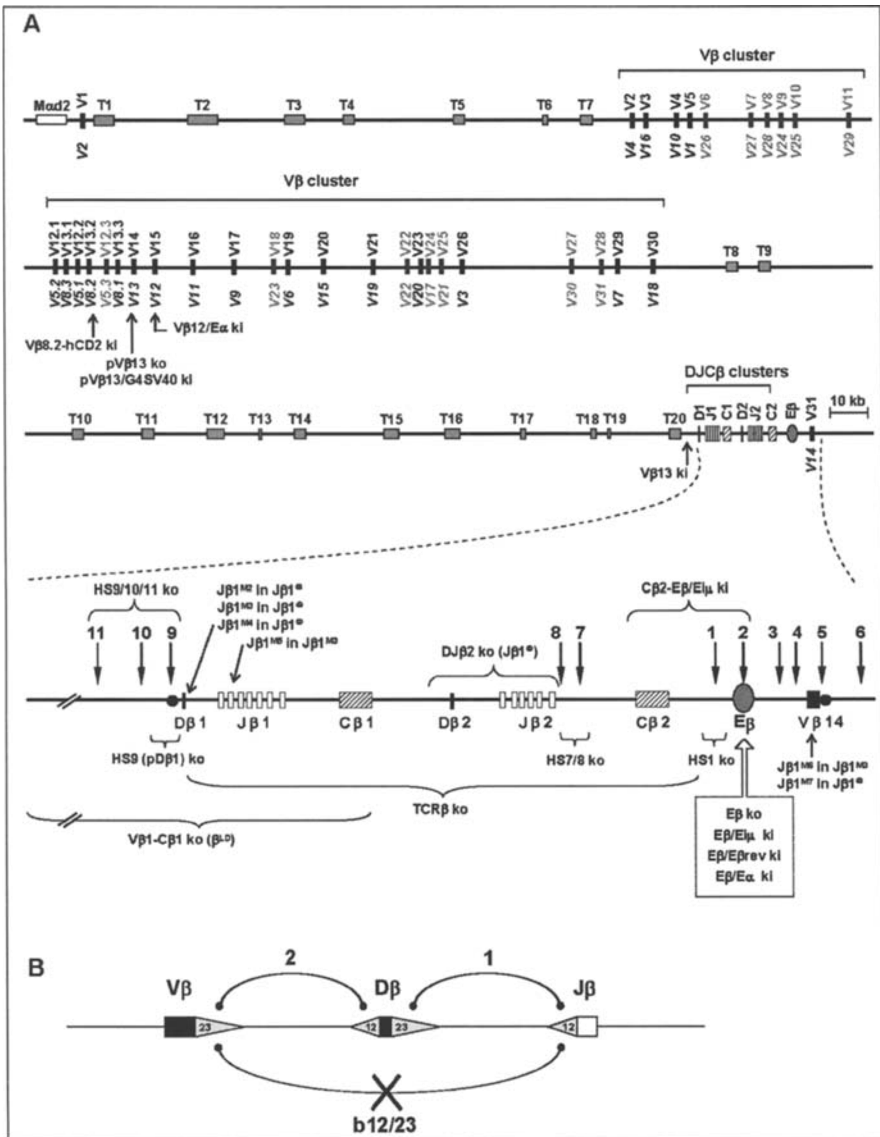


Figure 1. Structural organization and properties of the mouse *Tcrb* locus. A) The V, D, J and C segments are designed following the conventional ImMunoGeneTics (IMGT) nomenclature (top) and corresponding (published) names (bottom; for Vβ segments only) (see: http://imgt.cines.fr/textes/IMGTrepertoire/LocusGenes/nomenclatures/mouse/TRB/TRBV/Mu_TRBVnom.html). Thick black and grey lines represent functional and nonfunctional Vβ gene segments, respectively; and grey boxes represent trypsinogen genes. Thin black lines are for D gene segments; white and dashed boxes for J and C gene segments, respectively. An enlargement of the 3′ region is shown, where the Eβ enhancer (grey oval) and transcriptional Dβ1 and Vβ14 promoters (black ovals) are also figured. Published knock-out (ko) and knock-in (ki) alleles are indicated (see Table I for details and references). B) Schematic representation of *Tcrb* gene rearrangements. Dβ-to-Jβ joining occurs first (1) followed by Vβ-to-DJβ assembly (2). The beyond (b)12/23 constraint prohibits direct Vβ-to-Jβ joining. 12- and 23-RSSs are indicated.

paradoxical behaviors were initially analyzed using a modified *Tcrb* allele (Jβ1$^{\omega}$) containing the Dβ1-Jβ1 gene cluster only and derived mutants produced by successive gene-targeting (Fig. 1A and Table 1).[11] As such, using allelic mutants (Jβ1^{M3-M5}), it was shown that the 5'Dβ1-12RSS, but not the Jβ1-12RSS, efficiently targets Vβ-23RSSs for recombination, a phenomenon termed "beyond 12/23" (B12/23).[11] Moreover, 3'Dβ1-23RSS deleted alleles (Jβ1^{M2}) readily produced Vβ-to-Dβ rearrangements showing, at the very least, that these events do not require prior DJβ assembly.[12] Subsequently, it was shown that T-cells from mice in which the Vβ14-23RSS was replaced by the 3'Dβ1-23RSS and the Dβ gene segment deleted (Jβ1^{M6} allele), exhibited a dramatic increase in Vβ14 gene rearrangements, all directed to the Jβ-12RSSs as expected.[13] However, surprisingly, in the presence of an intact Dβ1 gene segment (Jβ1^{M7} allele), most were readily targeted to DJβ rearranged intermediates.[14] Collectively, these results demonstrate that RSSs can impose significant constraints on *Tcrb* gene assembly beyond enforcing the 12/23 rule. In particular the Dβ RSSs may function in an efficient manner to target the V(D)J recombinase, thus ensuring the utilization of a Dβ gene segment during variable region gene assembly and hence the diversity of the TCRβ repertoire.[15] However, the impact of this constraint(s) on regulatory events such as *Tcrb* ordered rearrangements remains largely unexplored.

Molecular Mechanism(s) of B12/23

Concurrent efforts by several laboratories have provided compelling evidence that B12/23 is established largely by the V(D)J recombinase and *Tcrb* RSSs and is enforced at or prior to DNA cleavage.[16-19] Additional work using hybrid RSSs has enlightened the role of the β-12-RSS nonamer and spacer sequences in imposing this constraint.[18-20] Finally, building on these earlier findings, the Schatz laboratory recently demonstrated the implementation of the B12/23 constraint via a low efficiency DNA cleavage reaction at various discrete stages dependent on the gene segments considered (e.g., initial single strand nicking at Jβ substrates; and, also, synapsis between Vβ and Jβ substrates).[21] Still the question remains as to how the 3'Dβ-23RSSs overcomes a possible 'wedge' in Jβ-12RSS cleavage to sustain an efficient Dβ-to-Jβ recombination. The high conservation (across mammalian phylogeny) of Dβ-RSSs compared to Jβ-12RSSs,[8] implying a strong evolutionary pressure on the former, may be relevant to the maintenance of this precise function.

Cis-Regulatory Elements at the *Tcrb* Locus

Comprehensive mapping of DNAse-I hypersensitive (HS) sites within the 3' side of the *Tcrb* locus has revealed 11 HS sites (named HS1 to HS11) encompassing a region from 3-kb upstream of Dβ1 to 3-kb downstream of Vβ14 (Fig. 1A).[22,23] Strikingly, HS2 and HS9 respectively overlap with Eβ and the promoter located upstream of Dβ1 (pDβ1),[24,25] both critical regulators of regional and more localized recombination events within the Dβ-Jβ clusters (see below). Similarly, HS5 overlaps with the Vβ14 promoter. This led to the assumption that other HSs also represent potential *cis*-regulatory elements for *Tcrb* gene rearrangements and/or transcription. To date, gene-targeting and the analysis of the resulting lymphoid cell phenotype have addressed the function of most of these putative recombinational/transcriptional elements, as well as that of at least one Vβ promoter (summarized in Fig. 1A and Table 1).

The transcriptional enhancer of the *Tcrb* locus (Eβ) has been described as a critical element in sustaining high level expression of a rearranged *Tcrb* gene.[26-28] Initial characterization of Eβ sequences has defined seven transcription-factor (TF) binding sites (named Eβ1 to Eβ7).[29,30] Efforts to further dissect the structural organization of Eβ and define minimal (core) enhancer sequences revealed the importance of a common composite (ETS/RUNX) TF motif, found within Eβ4 and Eβ6, in mediating enhancer activity.[31,32] In parallel, by a transgenic approach, a possible additional and important role of Eβ was shown in the control of *Tcrb* gene rearrangements.[33,34] Indeed, the generation of Eβ knock-out mice (Eβ$^{-/-}$) demonstrated a requirement of this element for *Tcrb* gene recombination and, at least in the Dβ-Jβ-Cβ regions, transcriptional expression (refs. 35,36; also see below). Accordingly, the Eβ$^{-/-}$ mice display a partial block of thymocyte differentiation at the DN cell stage and absence of αβT-cells in peripheral lymphoid organs,[37,38] a phenotype similar to

Table 1. Gene-targeting studies at the TCRβ locus

Genotype[1]	Description	Phenotype[2]	Refs.[3]
TCRβ ko	~15-kb deletion, from Jβ1.3 to Cβ2	Block in αβT-cell differentiation	39
Eβ ko	560-bp deletion of Eβ-containing sequences	Strongly reduced D-J and V-DJ rgts; reduced CA at DJCβ; Vβ regions not affected	35,36,38
Eβ/Eiμ ki	Eβ-to-Eiμ replacement	Dβ-Jβ and Vβ-DJβ rgts restricted to T-cells	36
Eβ/Eβrev ki	Inversion of Eβ	No phenotype	73
Cβ2-Eβ/Eiμ ki	11-kb replacement (from Cβ2 to Eβ) by Eμ	T-cell lineage restriction disrupted; Jβ1, but not Jβ2, rgts impaired	44
Eβ/Eα ki	Eβ-to-Eα replacement	Rgts still restricted to DN cells; CA impaired in DN, but not DP cells; less efficient TCRβ rgts	42
HS1 ko	780-bp deletion of HS1 (located 400-bp upstream of Eβ)	No phenotype	23
HS9/10/11 ko	3-kb deletion of sequences encompassing HS9, 10 and 11 (located upstream of Dβ1)	Decreased Dβ1 rgts; Dβ2 not affected; GL transcription of the Dβ1 region abolished	47
HS9 (pDβ1) ko	390-bp deletion removing the pDβ1 core region (HS9)	Decreased Dβ1 rgts; Dβ2 not affected; Reduced CA at the Dβ1Jβ1 region	48
HS7/8 ko	Deletion of the intronic sequences encompassing HS7 and HS8	No phenotype	46
Jβ1$^{\omega}$	Deletion of the Dβ2-Jβ2 region	No phenotype	11
Jβ1^{M2}	Mutation of the 3'Dβ1-RSS in the Jβ1$^{\omega}$ allele	Vβ-to-Dβ1 rgts are readily observed	12
Jβ1^{M3}	Deletion of Dβ1 and flanking RSS in the Jβ1$^{\omega}$ allele	~No Vβ rgts to any of the Jβ segments; Developmental block at the DN cell stage	11
Jβ1^{M4}	5'Dβ1-RSS to Jβ1.2-RSS replacement in the Jβ1$^{\omega}$ allele	Efficient DJβ rgts; Reduced levels of VDJβ rgts; Developmental block at the DN cell stage	11
Jβ1^{M5}	Jβ1.2-RSS to 5'Dβ1-RSS replacement in the Jβ1^{M3} allele	Vβ segments rearrange exclusively to the Jβ1.2-5'Dβ1-RSS; Normal numbers of αβT-cells	11
Jβ1^{M7}	Vβ14-RSS to 3'Dβ1-RSS replacement in the Jβ1$^{\omega}$ allele	Dramatic increase in Vβ14^{+} T-cells; Vβ14 rgts targeted to DJβ1, but not to Jβ1 (B12/23 rule not broken); Normal numbers of αβT-cells	14

continued on next page

Table 1. Continued

Genotype[1]	Description	Phenotype[2]	Refs.[3]
Jβ1^{M6}	Vβ14-RSS to 3'Dβ1-RSS replacement in the Jβ1^{M3} allele	Dramatic increase in Vβ14$^+$ T-cells; Direct Vβ14 to Jβ1 rgts Normal numbers of αβT-cells; AE maintained	13
Vβ1-Cβ1 ko (β^{LD})	475-kb deletion of *Tcrb* sequences from Vβ1 to 3'Cβ1	Increased rgts using the most proximal Vβ gene (Vβ10); Increased CA at Vβ10; AE maintained; Aberrant Vβ10-Dβ2 and Vβ10-Jβ2 rgts are observed	102
Vβ8.2/hCD2 ki	Insertion of a human CD2 cDNA downstream of Vβ8.2	Bi-allelic expression of hCD2 prior to V-DJ rgts Distance-dependent expression of hCD2 in VDJ rearranged alleles	81
Vβ12/Eα ki	Eα insertion upstream of Vβ12	Increased rgts using Vβ12: AE partially subverted, but feedback inhibition is maintained	79
Vβ13 ki	Vβ13 gene segment and promoter inserted 6.8-kb upstream of Dβ1	Same utilization frequency of the inserted and endogenous Vβ13 copies; AE subverted at the level of rgts, but not at the level of transcription	53
pVβ13 ko	Deletion of a 1.2-kb region containing the Vβ13 promoter	Vβ13 rgts inhibited and CA reduced; AE maintained	52
pVβ13/G4SV40 ki	1.2-kb replacement of pVβ13 by the SV40 minimal promoter +5 copies of Gal4 binding sequences	Normal levels of Vβ13 rgts; Cleavage at Vβ13 not significantly affected (but mostly abnormal regarding the cleavage site); TCRβ Tg blocks Vβ13 rgts, but not the cleavages at the aberrant sites	52

[1]ko: knock-out mouse; ki: knock-in mouse; rgts: rearrangements; AE: allelic exclusion; CA: chromatin accessibility.
[2]Additional phenotypes are described in the text.
[3]Additional articles describing the phenotype(s) at the targeted alleles are mentioned in the text.

the one observed in *Tcrb* knock-out mice.[39] Of note, besides these impacts on *Tcrb* gene expression, recombination and αβT-cell development, further analyses provided evidence for an accumulation of rare recombination products in Eβ-deleted T-cells, including coding joints between the Dβ1 and Dβ2 gene segments and intermediate products of RSS cleavage [so called signal ends (SE)] at Vβ14 and Dβ gene segments (in DP cells).[38,40,41] The implication(s) of these observations regarding a physiological role of Eβ in the correct pairing of Dβ-Jβ gene segments and/or the processing of their recombination products still remains unclear.

Results from transgenic mouse experiments generally argue for a role of lymphoid enhancers in mediating tissue- and stage-specific induction of antigen-receptor gene recombination.[1] Importantly, this does not necessarily imply that the patterning of endogenous gene rearrangements be simply regulated by enhancer sequences alone. For example, whereas *Tcra* enhancer

(Eα)-induced rearrangements within a *Tcrb* transgenic substrate occurred efficiently at the DP cell stage, Eβ-to-Eα replacement within the endogenous *Tcrb* locus still resulted in DN-restricted recombination, although at a low efficiency.[33,42] Likewise, *Tcrb* transgenic substrates bearing the *Igh* intronic enhancer (Eiμ) demonstrated Dβ-to-Jβ rearrangements in both T and B-cells and Vβ-to-DJβ rearrangements in T-cells only.[43] However, on replacing Eβ for Eiμ at the endogenous locus, recombination was then restricted to T-cells and Dβ-to-Jβ joining.[36] Finally, knock-in replacement of Eβ plus Cβ2-Eβ intervening sequences by Eiμ yielded significant levels of Dβ-to-Jβ and Vβ-to-DJβ rearrangements in T and B-cells, unlike the mere Eβ-to-Eiμ replacement just mentioned.[44] Altogether, these data strongly suggest that additional *cis*-regulatory elements may contribute to the control of lineage- and developmental stage-specificity of V(D)J rearrangement events at the *Tcrb* locus. In this context, deletion of either the Eβ-proximal HS1 or HS7-HS8 sites resulted in no obvious phenotype.[23,45,46] Therefore, if functionally relevant, these elements may be expected to display high levels of redundancy. Testing this hypothesis will require further gene-targeting efforts to combine discrete deletions within the same allele.

Additional *cis*-elements shown to play a role in regulating recombination include the Dβ1- and Vβ13-associated promoters. T-lymphocytes from knock-out mice in which pDβ1 (either alone or together with the upstream HS10 and HS11 sites) has been deleted, exhibited a specific block in recombination of the Dβ1 gene segment whereas the Dβ2 gene segment rearranged normally.[47,48] As germline expression of the Dβ2 segment likely depends on an as yet uncharacterized regional promoter, common opinion states that Dβ2 recombination also relies on such a promoter-mediated, localized control, though this has yet to be formally demonstrated. Such reasoning was extended to the control of rearrangement of individual Vβ gene segments, as each is preceded by a transcriptional promoter, only a few of which have been thoroughly characterized.[49-51] Indeed, a 1.2-kb deletion of sequences upstream of Vβ13 resulted in the inhibition of transcription of the corresponding V-gene segment and decreased RAG1/2-mediated rearrangement.[52] Moreover, when inserted ~6-kb upstream of Dβ1, a supplemental Vβ13 gene segment (and associated promoter) displayed the same frequency of rearrangement as the Vβ13 endogenous homologue.[53] These studies support the hypothesis that Vβ promoters may generally be sufficient to induce recombination of their associated V-gene segments. Nonetheless, the presence of a global regulator within the *Tcrb* locus and more particularly the 5' Vβ region, notably controlling long-range *cis*-interactions between the Vβ and Dβ-Jβ-Cβ clusters (see below) cannot be ruled out.

Trans-Regulators of *Tcrb* Locus Expression/Recombination

The critical role of transcriptional *cis*-regulatory elements in orchestrating V(D)J rearrangements may be clearly linked to (at least some of) the TFs they normally recruit. Initially, TF binding motifs (Eβ1-Eβ7) of Eβ were analyzed using dedicated bandshift and/or in vitro footprinting as well as gene-reporter assays.[2,30] The identified motifs included those for TFs belonging to the GATA, ATF/CREB, bHLH, ETS and RUNX families. Subsequently, most of these sites appeared readily occupied in developing T-cells (according to in vivo footprinting assays; ref. 32 and our unpublished data); and chromatin immunoprecipitation (ChIP) experiments further demonstrated the specific binding of TFs RUNX1 and ETS1 to Eβ overlapping sequences.[54,55] Concordantly, overexpression of *Runx1* in an ex-vivo T-cell differentiation model system resulted in increased levels of *Tcrb* expression.[56] Likewise, the combination of in vivo genomic footprinting and ChIP assays has implied the loading of a battery of TFs to pDβ1-overlapping sites, including Ikaros, ETS1, RUNX1, ATF/CREB, GATA, SP1 and KLF5.[54,57,58] Although not yet completely characterized, promoter regions 5' of Vβ genes generally appeared to be also enriched with a variety of TF-binding sites, including the ETS/RUNX and ATF/CREB composite motifs.[3,50,51]

Due to the pleiotropic effects that these factors generally exert throughout embryonic and/or T-cell development (reviewed by Rothenberg et al[59]), a definitive demonstration of their genuine role in the control of gene transcription and/or rearrangement at the *Tcrb* locus represents a difficult challenge. To date, five nuclear factors/signaling pathways have been reported as directly interfering with *Tcrb* gene rearrangements.[2] These include TFs HEB,[60] c-MYB,[61] and BCL11b;[62]

the scaffold/matrix-associated region 1 (SMAR1) protein;[63] and, probably, transcriptional effectors downstream of the Notch1 signaling cascade.[64] Remarkably, disruption of most of the corresponding genes (and overexpression of SMAR1), resulted in a selective impairment of Vβ-to-DJβ recombination,[60,61,63,64] underlining the uniqueness of this particular step of the *Tcrb* variable region gene assembly. Although not a trivial task, the complementary approach of mutating the cognate binding motif(s) within *Tcrb* endogenous *cis*-regulatory elements may in the end be required to firmly ascertain the direct contribution of each and every factor to this process.

Chromatin Accessibility

Long believed to be solely confined to a static role as a DNA packaging envelope, chromatin is now also recognized as a master player in the dynamic regulation of all genomic DNA transactions. Indeed, at the command of developmental/signaling inputs, chromatin evolves from a hermetically packed (heterochromatin) shell to a more relaxed, 'metabolism-amenable' (euchromatin) structure. In this context, from the widespread correlation between ongoing recombination and transcription through nonrearranged gene segments/regions that prevails at antigen receptor loci, to the demonstration of a critical impact of both *Ig*/*Tcr* transcriptional *cis*- and *trans*-regulators on the former activity, all arguments converged to the model where lineage- and/or temporal-specific actions of a common V(D)J recombinase are primarily regulated at the level of chromatin permissivity.[65] The formal demonstration that these controls indeed depend on the regulated changes in chromatin accessibility came from experiments using isolated nuclei and purified RAG1/2 factors.[66] Within a given locus, chromatin, associated with V(D)J rearranging gene segments, generally adopts a 'less-compacted' configuration compared to that spreading over recombination inert regions (for review see refs. 67,68). As for the *Tcrb* locus, several molecular parameters synonymous with euchromatin (germline transcription; lack of CpG methylation; enrichment in histone H3/H4 acetylation and H3K4 methylation; accessibility to restriction enzymes) have been correlated with stage-specific (DN2/DN3) V(D)J recombination events.[23,38,41,42,48,57,69-71] Conversely, from the DN3-to-DP thymic-cell transition (i.e., β-selection) onwards, heterochromatin features were generally found along the chromosomal regions comprised of nonrearranged Vβ genes.[23,41,70,72]

Chromatin Remodeling by Enhancer-Promoter Interaction

The stringent T-cell phenotype observed in the $E\beta^{-/-}$ mouse made it an excellent model system to investigate the role of transcriptional enhancers in regulating recombinational accessibility. These studies were significantly helped by the possibility of analyzing chromatin-associated parameters in Eβ-deleted, early-developing T-cells unable to proceed with V(D)J recombination due to RAG deficiency ($E\beta^{-/-}$ x $Rag^{-/-}$ mice). Indeed, detailed analysis of DN T-cells from these animals and comparison with those from $Rag^{-/-}$ controls, provided compelling evidence that Eβ mediates chromatin remodeling within the proximal Dβ-Jβ-Cβ domains and, likely, activates the germline promoters flanking the Dβ gene segments (Fig. 2).[54,57,69] In sharp contrast, the unaffected appearance of the chromatin structure at Eβ-deleted alleles on the distal 5' Vβ and neighboring 3' Vβ14 regions suggested no impinging effect of Eβ at these particular locations. Despite such a polarized enhancer activity, however, inversion of Eβ sequence orientation did not alter *Tcrb* gene regulation (in terms of transcription and recombination), thus supporting a DNA-looping rather than linking/tracking mechanism for this enhancer's role in the *cis*-activation of target promoter/recombination sequences.[73] In parallel, experimental approaches analyzing the chromatin remodeling function of the Dβ1 and Vβ13 promoters, showed that both act to reduce local chromatin accessibility, i.e., without affecting the neighboring Dβ2-Jβ2 or Vβ gene segments, respectively (Fig. 2).[48,52,54,74]

Thorough molecular analysis of Eβ-deleted thymocytes subsequently demonstrated that this element contributes to the assembly of a functional nucleoprotein complex at pDβ1, including the loading of discrete TFs and basal transcriptional machinery.[57] Such a dedicated process likely involves a physical interaction between the two *cis*-regulatory elements, possibly contributing to the formation of a stable holocomplex (Fig. 3).[54] In this context, the chromatin accessibility of Jβ1 gene segments is relatively unaffected at pDβ1-deleted alleles, implying that Eβ may exert both

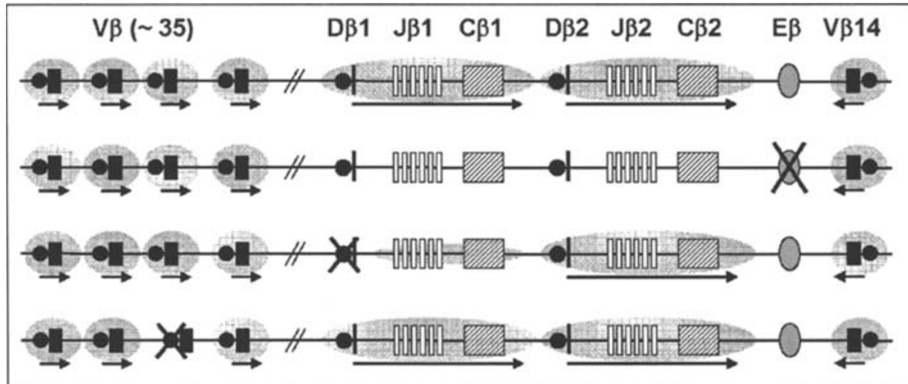


Figure 2. Effects on chromatin accessibility of mouse knock-out deletion of *Tcrb cis*-regulatory elements. The upper line shows a schematic representation of the *Tcrb* locus in wild-type mice. The lower (2nd, 3rd and 4th) lines summarize the effects on chromatin accessibility of the Eβ, pDβ1 and pVβ targeted deletions, respectively. Shadowed areas symbolize relaxed, accessible chromatin; arrows indicate germline transcription.

pDβ1-dependent and -independent chromatin remodeling functions.[54] Overall, these results support a model whereby the two elements act in coordination to regulate chromatin accessibility at the proximal Dβ-Jβ-Cβ regions.

Chromatin Remodeling Enzymes and the Control of Tcrb Locus Expression/ Recombination

Chromatin structure can be altered in a number of different ways, including covalent modifications (e.g., acetylation, methylation, phosphorylation and ubiquitinylation) of histone tails, replacement by distinct histone variants and changes in nucleosome positioning via the action of specialized enzymatic and/or chromatin remodeling activities.[75] A synopsis of these activities and dedicated factors potentially involved in regulating V(D)J recombinational accessibility—extending beyond the scope of this chapter—has recently been proffered.[68] Regarding *Tcrb* D-J gene segments for instance, the likelihood that such factors directly interfere with their chromosomal accessibility was brought to light in experiments where targeted recruitment to pDβ1 of the H3K9 specific histone methyltransferase G9a (using protein fusion to the DNA-binding domain of Gal4 and a modified *Tcrb* minilocus transgene) was shown to induce an extensive change in local chromatin environment (i.e., towards an heterochromatin-like layout) and a significant decrease in the onset of Eβ-mediated Dβ-Jβ transgenic expression and recombination.[76] Thus far, evidence indeed exists for a few euchromatin inducers having an impact on the onset of V(D)J recombination at the *Tcrb* endogenous locus.[67,68] Using ChIP assays, we and others have shown that BRG1 (a subunit of the nucleosome-disrupting complex SWI/SNF) and the histone acetylases CBP/p300 and PCAF, are associated with Eβ- and/ or pDβ1-overlapping sequences in primary DN thymocytes and/or pre-T-cell lines.[54,57,71] Moreover, Gal4-mediated targeting of BRG1 to a *Tcrb* transgenic substrate completely substitutes the pDβ1 function in inducing local recombinational accessibility; and knock-down by RNA interference of two essential SWI/SNF components (BRG1 and BRM) results in decreased accessibility of endogenous Dβ-RSSs.[77] As suggested by these authors, formation of an Eβ/pDβ1-based holocomplex may generate a new interaction surface for the stable association of SWI/SNF components, which would then contribute to remodel or displace a neighboring nucleosome(s), thus exposing sites required for the initiation of germline transcription and/or recombination. Consistent with this hypothesis, Dβ-RSSs and immediate surrounding sequences display high nucleosome densities[57] and may comprise strong nucleosome-positioning sites.[78] Whether additional chromatin regulators play a similar, complementary, or differing role(s) in controlling *Tcrb* locus accessibility remains to be determined.

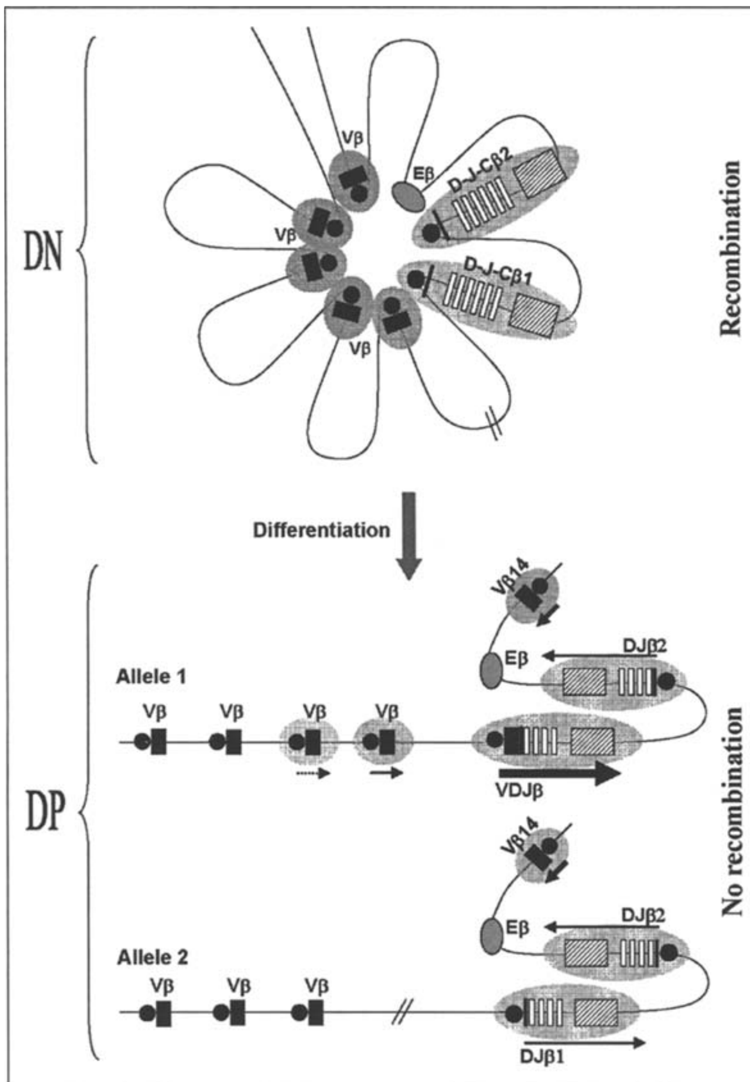


Figure 3. Chromosomal contraction/expansion events at the *Tcrb* locus during early T-cell development and prospective impact on the control of V(D)J recombination and allelic exclusion. Legends for the gene segments, regulatory elements and locus organization are as in Figure 2. In DN thymocytes, chromosomal contraction and DNA looping bring the various *cis*-regulatory elements within the *Tcrb* locus closer, enabling potential interaction(s) and the onset of V(D)J *cis*-recombination. Differentiation to the DP cell stage results in locus decontraction and heterochromatinization of most unrearranged Vβ gene segments, thus preventing further Vβ-to-DJβ rearrangement at these sites. At this stage, the VβDJβ rearranged variable gene region, under the control of Eβ, is highly transcribed (Allele 1). However, unrearranged Vβ gene segments located immediately upstream of the VβDJβ rearranged region (Allele 1) and the Vβ14 gene segment (Alleles 1 and 2), remain accessible, yet do not rearrange. For didactic purposes, the allelic conformations represented here were chosen to result from (i) a productive rearrangement on the first Vβ-to-DJβ joining attempt; (ii) a VβDJβ rearranged variable gene region made of a 5', middle-located Vβ gene segment joined to Dβ1/Jβ1 gene segments.

Beyond Chromatin Accessibility

The B12/23 constraint discussed above alone provided evidence that the regulation of *Tcrb* V(D)J recombination goes far beyond the already sophisticated process of an appropriate spatial and temporal tuning of chromosomal access to the particular RSSs. In addition however, studies surrounding the inhibition of *Tcrb* gene rearrangements once the developing thymocytes have passed through β-selection and reached the DP thymic cell stage (where rearranged *Tcrb* gene expression is maintained and V(D)J recombination targeted towards the *Tcr-Ja* locus), have recently revealed puzzling results on this matter. Firstly, the few $E\beta^{-/-}$ thymocytes which differentiate into DP cells (via transient expression of a γδ TCR—see ref. 37), partially recover chromatin accessibility over their Dβ-Jβ-Cβ (and Vβ14) domains, indicating no further requirement of Eβ to unwrap heterochromatin over these domains at this stage.[38,41] Yet Vβ14-to-DJβ rearrangements are still not observed at these sites. Secondly, in mice in which Eβ was replaced by Eα, or in which Eα was inserted downstream of a Vβ gene segment in the 5' Vβ domain, *Tcrb* gene rearrangement was still inhibited in DP cells despite signs of chromosomal accessibility within the region adjacent to the replaced/newly inserted enhancer.[42,79] Thirdly, Vβ gene segments located 5' (up to 150-kb) of a rearranged VDJβ unit are maintained in a relaxed chromosomal form, yet remain refractory to V(D)J recombination.[80,81] It appears that once a functional variable gene region has been made, further *Tcrb* gene rearrangement is inhibited via an (epigenetic?) process(es) acting relatively independently of mere chromatin accessibility.

What are the molecular mechanism(s) that enforce the inhibition of *Tcrb* gene recombination in DP thymocytes? Pioneering studies using fluorescence in situ hybridization (FISH) have revealed allele subnuclear (re)positioning and large-scale locus contraction/ chromosomal looping as novel processes that may also be involved in the developmental regulation of gene expression and recombination at immune loci.[67,82] Recently, using both FISH and chromosome conformation capture (3C) assays, the *Tcrb* and *Tcra* loci were indeed shown to undergo long-range DNA contraction in DN and DP thymocytes, respectively.[83] The folding of the *Tcrb* locus is reversed at the DP stage, raising the intriguing possibility that this locus contraction/expansion process contributes to regulating both Vβ-to-DJβ recombination in early DN cells and its inhibition in more developed DP cells, respectively (Fig. 3). How *Ig*/ *Tcr* locus contraction/expansion is achieved is still unclear. Deciphering the underlying molecular mechanism(s) and factors involved will significantly improve our understanding of long-range synapsis events in V(D)J recombination and their regulation.

Allelic Exclusion at the *Tcrb* Locus

Individual lymphocytes generally synthesize antigen receptors of a unique specificity. Accordingly, the vast majority of mature αβT-cells bear a single TCRβ chain (out of a possible two, one encoded at each *Tcrb* allele).[84-86] This is in part achieved by a phenomenon referred to as allelic exclusion.[87] In fact, similar to the situation first described at the *Igh* and *Igk* loci in B-cells,[88] approximately 60% of αβT-cells harbor a single productive (in frame with Cβ) VDJβ rearranged variable region gene, whereas the remaining 40% carry two VDJβs rearranged in nonproductive and productive configurations, respectively. The 60/40 ratio is consistent with one productive in every three rearrangements at individual V-to-DJ joints and a feedback signal whereby a TCR/Ig product from one functionally rearranged allele leads to the inhibition of further V-to-DJ joining on the opposite allele. This regulated model of allelic exclusion clearly implies an initiation step whereby homologous alleles are sequentially targeted for recombination.[87]

Initiation of Allelic Exclusion at the Tcrb Locus

Studies mainly carried out at *Igh* and *Igk* loci have led to two distinct types of mechanistic model explaining how V(D)J recombination may be initiated in an allelic-specific manner.[82,87] On the one hand, instructive models call upon epigenetics [i.e., DNA (de)methylation, histone tagging, nuclear (pericentromeric) positioning and/or other less-well defined epigenetic mark(s) shown by asynchronous DNA replication] to differentially label the two alleles such that only one will

be prone to rearrange. On the other hand, stochastic models evoke inter-allelic competition (and generally a low probability in allelic activation) as a means to dissociate allele rearrangements. In this context, analyses at the *Tcrb* locus brought contrasting results.[2,89] Indeed, several findings argued against an intrinsically favored recombinational accessibility towards one allele only at this locus. When investigated, DNA demethylation at the Dβ-Jβ-Cβ regions and germline transcription of a Vβ-containing region appeared, at least initially, biallelic (refs. 23,69,81 and our unpublished results). In addition, Vβ-to-DJβ recombination has been shown to initiate on one allele before completion of all possible VDJβ rearrangements on the opposite allele.[90] Nonetheless, like *Ig* alleles, *Tcrb* alleles seem to replicate asynchronously in developing thymocytes, with one often being recruited to pericentromeric heterochromatin.[83,91] Additional efforts will thus be required to reconcile these apparently conflicting observations.

Feedback Inhibition of Tcrb Recombination

In late DN3 thymocytes, the assembly of the newly formed TCRβ chain with the invariant preTα chain and CD3 complex subunits forms the pre-TCR (reviewed by von Boehmer et al[92]). Pre-TCR-mediated signaling—which in addition to the pre-TCR components involves many downstream kinases and adaptator molecules such as e.g., p56[lck] and SLP-76—ensures β-selection and a number of critical outcomes for cells bypassing this checkpoint (i.e., the maintenance of cell survival, induction of cell proliferation and differentiation into DP thymocytes and allelic exclusion). As a result, disruption of pre-TCR signaling by gene inactivation of *pTa*, *cd3* or *slp-76*, blocks αβT-cell development at the DN thymic cell stage and impairs allelic exclusion.[93-95] Conversely, enforced expression of a *Tcrb* transgene in DN thymocytes inhibits endogenous Vβ-to-DJβ rearrangements and promotes developmental progression into DP cells.[96,97] A variety of basic processes have been suggested to account for the suppression of *Tcrb* gene rearrangements by pre-TCR signaling, including cell-cycle-dependent degradation of RAG2 and Vβ gene silencing.[92] In addition, as described below, a dissection of pre-TCR downstream signaling demonstrated that immature T-cells utilize distinct pathways to achieve allelic exclusion versus cell expansion and differentiation.[2,89]

The pre-TCR promotes activation of multiple signaling pathways including Ca^{2+} flux, protein kinase C (PKC) and RAS-RAF-MAP kinase (MAPK) signaling pathways. Activation of pre-TCR proximal p56[lck] or more distal PKC kinases is sufficient to induce all aspects of β-selection, including allelic exclusion.[88,98] Strikingly however, in DN thymocytes, small GTPases RAS- or RAF-mediated activation of the MAPK pathway induces T-cell expansion and cellular differentiation but does not block *Tcrb* gene rearrangement, implying a partition of signal transduction for the feedback inhibition of Vβ-to-DJβ joining and for cellular expansion/differentiation somewhere between the PKC and RAS/RAF activation nodes.[88,99] However, a normal heterochromatin layout is observed along the Vβ locus in DP thymocytes generated via MAPK activation,[100] indicating distinct requirements in DP and DN cells to sustain inhibitory features. Reduced accessibility of Vβ gene segments in DP thymocytes likely contributes to lock out allelic exclusion,[41,70,72] relying on the setting of an appropriate developmental program via the induction of discrete TFs (reviewed by Rothenberg et al[59]). Concordantly, notably, *Ets1*-deficient thymocytes were shown to display a disruption of allelic exclusion and impaired DN3-to-DP cell transition.[101] An ultimate goal will be to link pre-TCR signal transduction cascades with all the nuclear functions involved in securing allelic exclusion, including those particular factors that possibly reduce chromatin accessibility at Vβ promoters thereby repressing transcription and recombination.

Allelic exclusion likely involves multiple layers of control in addition to the mere changes in chromatin accessibility.[2] Indeed, inhibition of Vβ gene rearrangement in DP thymocytes is preserved when Vβ gene segments are maintained in a transcriptionally active (accessible) configuration by insertional knock-in of Eα.[42,79] In this context, as mentioned, 'taking away' Vβ genes by pericentromeric allele recruitment or *Tcrb* locus decontraction could play a significant role.[83] In support of this latter possibility, distinct gene knock-in experiments, in which a Vβ gene was introduced in proximity (5') to the Dβ1 or Dβ2 region, demonstrated an escape from allelic

exclusion (at the level of Vβ-to-DJβ or Vβ-to-Dβ recombination, respectively).[53,102,103] However, these mechanisms still cannot explain the maintenance of an inhibition for recombination of the unrearranged (and apparently accessible) Vβ segments located 5' proximal (in *cis*) of a rearranged VβDJβ variable gene region.[80,81] Hence, supplementary constraints may act in an as-yet undefined way to eventually enforce allelic exclusion.

Conclusion and Future Direction

Because the immune system is not absolutely required for survival in a pathogen-free environment, antigen receptor genes have served as tractable models to study the regulation of gene expression at complex genomic loci using most notably gene targeting technologies. As has been illustrated here, the *Tcrb* locus is one of the most genetically modified in the mouse so far. During the past decade, these studies have led to a better understanding of the function of distinct *cis*-regulatory elements and their hierarchical impact in the control of locus expression and recombination in terms of chromatin accessibility (Fig. 2 and 3). Moreover, previously unsuspected processes have been unraveled, including subnuclear localization and locus contraction/expansion, not to mention the RSS biased usage. It remains to be resolved which processes are of general use during (lymphoid-) cell differentiation programs and which are more specific to the control of V(D)J recombination, or even *Tcrb* locus expression.

Additional findings will undoubtedly complete our knowledge in the near future. In particular, progresses should include a better characterization of the nucleoprotein complexes recruited to *Tcrb*-regulatory elements, the resulting epigenetic features orchestrating accessibility/heterochromatinization at this locus and plausible interplays with DNA-modifying machineries [including the RAG2 and/or additional component of the V(D)J recombinase; refs. 104-106]. The recent development of large-scale, genome-wide ChIP-on-chip methodologies looks particularly promising in this regard. In parallel, it will be important to establish whether intergenic and/or antisense transcripts (see chapters by M. Krangel and A. Corcoran) are also produced at the *Tcrb* locus and, if so, investigating what impact they have on accessibility to the V(D)J recombinase. Likewise, does 'transcriptional interference' play a role at this locus? Finally, a better understanding of the precise mechanisms leading to the establishment and enforcement of allelic exclusion is still expected and, further still are insights into the potential risks for the immune system/organism caused by relaxed accessibility/allelic exclusion during *Tcrb* locus recombination.

Acknowledgements

We thank members of the PF lab for critical reading of this manuscript. Work in this lab is supported by institutional grants from Inserm and the CNRS and by specific grants from the 'Fondation Princesse Grace de Monaco', the 'Association pour la Recherche sur le Cancer' (ARC 3275 & 3136), the 'Agence Nationale de la Recherche' (IMPB100-045632 & ANR-BLAN-0396-03/BYOS-0006-01) and the Commission of the European Communities (MRTN-CT-035733). The Eβ-targeted mice studied in the PF lab were generated with the help of the functional genomic platform of the Marseille-Nice Genopole®.

References

1. Jung D, Alt FW. Unraveling V(D)J recombination; insights into gene regulation. Cell 2004; 116:299-311.
2. Jackson AM, Krangel MS. Turning T-cell receptor β recombination on and off: more questions than answers. Immunol Rev 2006; 209:129-141.
3. Glusman G, Rowen L, Lee I et al. Comparative genomics of the human and mouse T-cell receptor loci. Immunity 2001; 15:337-349.
4. Cauwelier B, Dastugue N, Cools J et al. Molecular cytogenetic study of 126 unselected T-ALL cases reveals high incidence of TCRβ locus rearrangements and putative new T-cell oncogenes. Leukemia 2006; 20:1238-1244.
5. Schlissel MS, Kaffer CR, Curry JD. Leukemia and lymphoma: a cost of doing business for adaptive immunity. Genes Dev 2006; 20:1539-1544.
6. Nemazee D. Receptor selection in B and T-lymphocytes. Annu Rev Immunol 2000; 18:19-51.

7. Ramsden DA, Baetz K, Wu GE. Conservation of sequence in recombination signal sequence spacers. Nucleic Acids Res 1994; 22:1785-1796.
8. Tillman RE, Wooley AL, Hughes MM et al. Regulation of T-cell receptor β-chain gene assembly by recombination signals: the beyond 12/23 restriction. Immunol Rev 2004; 200:36-43.
9. Livak F, Burtrum DB, Rowen L et al. Genetic modulation of T-cell receptor gene segment usage during somatic recombination. J Exp Med 2000; 192:1191-1196.
10. Olaru A, Patterson DN, Cai H et al. Recombination signal sequence variations and the mechanism of patterned T-cell receptor-β locus rearrangement. Mol Immunol 2004; 40:1189-1201.
11. Bassing CH, Alt FW, Hughes MM et al. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. Nature 2000; 405:583-586.
12. Sleckman BP, Bassing CH, Hughes MM et al. Mechanisms that direct ordered assembly of T-cell receptor β locus V, D and J gene segments. Proc Natl Acad Sci USA 2000; 97:7975-7980.
13. Wu C, Ranganath S, Gleason M et al. Restriction of endogenous T-cell antigen receptor β rearrangements to Vβ14 through selective recombination signal sequence modifications. Proc Natl Acad Sci USA 2007; 104:4002-4007.
14. Wu C, Bassing CH, Jung D et al. Dramatically increased rearrangement and peripheral representation of Vβ14 driven by the 3'Dβ1 recombination signal sequence. Immunity 2003; 18:75-85.
15. Hughes MM, Yassai M, Sedy JR et al. T-cell receptor CDR3 loop length repertoire is determined primarily by features of the V(D)J recombination reaction. Eur J Immunol 2003; 33:1568-1575.
16. Tillman RE, Wooley AL, Hughes MM et al. Restrictions limiting the generation of DNA double strand breaks during chromosomal V(D)J recombination. J Exp Med 2002; 195:309-316.
17. Olaru A, Patterson DN, Villey I et al. DNA-Rag protein interactions in the control of selective D gene utilization in the TCRβ locus. J Immunol 2003; 171:3605-3611.
18. Jung D, Bassing CH, Fugmann SD et al. Extrachromosomal recombination substrates recapitulate beyond 12/23 restricted V(D)J recombination in nonlymphoid cells. Immunity 2003; 18:65-74.
19. Tillman RE, Wooley AL, Khor B et al. Targeting of Vβ to Dβ rearrangement by RSSs can be mediated by the V(D)J recombinase in the absence of additional lymphoid-specific factors. J Immunol 2003; 170:5-9.
20. Hughes MM, Tillman RE, Wehrly TD et al. The B12/23 restriction is critically dependent on recombination signal nonamer and spacer sequences. J Immunol 2003; 171:6604-6610.
21. Drejer Teel AH, Fugmann SD, Schatz DG. The beyond 12/23 restriction is imposed at the nicking and pairing steps of DNA cleavage during V(D)J recombination. Mol Cell Biol 2007; 27:6288-6299.
22. Hashimoto Y. T-cell receptor β gene has two downstream DNase I hypersensitive regions. J Exp Med 1989; 169:2097-2107.
23. Chattopadhyay S, Whitehurst CE, Schwenk F et al. Biochemical and functional analyses of chromatin changes at the TCR-β gene locus during $CD4^-CD8^-$ to $CD4^+CD8^+$ thymocyte differentiation. J Immunol 1998; 160:1256-1267.
24. Sikes ML, Gomez RJ, Song J et al. A developmental stage-specific promoter directs germline transcription of DβJβ gene segments in precursor T-lymphocytes. J Immunol 1998; 161:1399-1405.
25. Doty RT, Xia D, Nguyen SP et al. Promoter element for transcription of unrearranged T-cell receptor β-chain gene in pro-T-cells. Blood 1999; 93:3017-3025.
26. Krimpenfort P, De Jong R, Uematsu Y et al. Transcription of T-cell receptor β-chain genes is controlled by a downstream regulatory element. EMBO J 1988; 7:745-750.
27. McDougall S, Peterson CL, Calame K. A transcription enhancer 3' of Cβ2 in the T-cell receptor β locus. Science 1988; 241:205-208.
28. Busse CE, Krotkova A, Eichmann K. The TCRβ enhancer is dispensable for the expression of rearranged TCRβ genes in thymic DN2/DN3 populations but not at later stages. J Immunol 2005; 175:3067-3074.
29. Takeda J, Cheng A, Mauxion F et al. Functional analysis of the murine T-cell receptor β enhancer and characteristics of its DNA-binding proteins. Mol Cell Biol 1990; 10:5027-5035.
30. Leiden JM, Thompson CB. Transcriptional regulation of T-cell genes during T-cell development. Cur Op Immunol 1994; 6:231-237.
31. Carvajal IM, Sen R. Functional analysis of the murine TCR β-chain gene enhancer. J Immunol 2000; 164:6332-6339.
32. Tripathi RK, Mathieu N, Spicuglia S et al. Definition of a T-cell receptor β gene core enhancer of V(D)J recombination by transgenic mapping. Mol Cell Biol 2000; 20:42-53.
33. Capone M, Watrin F, Fernex C et al. TCRβ and TCRα gene enhancers confer tissue- and stage-specificity on V(D)J recombination events. EMBO J 1993; 12:4335-4346.
34. Okada A, Mendelsohn M, Alt FW. Differential activation of transcription versus recombination of transgenic T-cell receptor β variable region gene segments in B and T lineage cells. J Exp Med 1994; 180:261-272.

35. Bouvier G, Watrin F, Naspetti M et al. Deletion of the mouse T-cell receptor β gene enhancer blocks αβ T-cell development. Proc Natl Acad Sci USA 1996; 93:7877-7881.
36. Bories JC, Demengeot J, Davidson L et al. Gene-targeted deletion and replacement mutations of the T-cell receptor β-chain enhancer: the role of enhancer elements in controlling V(D)J recombinational accessibility. Proc Natl Acad Sci USA 1996; 93:7871-7876.
37. Leduc I, Hempel WM, Mathieu N et al. T-cell development in TCRβ enhancer-deleted mice: implications for αβ T-cell lineage commitment and differentiation. J Immunol 2000; 165:1364-1373.
38. Ryu CJ, Haines BB, Draganov DD et al. The T-cell receptor β enhancer promotes access and pairing of Dβ and Jβ gene segments during V(D)J recombination. Proc Natl Acad Sci USA 2003; 100:13465-13470.
39. Mombaerts P, Clarke AR, Hooper M et al. Creation of a large genomic deletion at the T-cell antigen receptor β-susunit locus in mouse embryonic stem cells by gene targeting. Proc Natl Acad Sci USA 1991; 88:3084-3087.
40. Hempel WM, Stanhope-Baker P, Mathieu N et al. Enhancer control of V(D)J recombination at the TCRβ locus: differential effects on DNA cleavage and joining. Genes Dev 1998; 12:2305-2317.
41. Mathieu N, Spicuglia S, Gorbatch S et al. Assessing the role of the T-cell receptor β gene enhancer in regulating coding joint formation during V(D)J recombination. J Biol Chem 2003; 278:18101-18109.
42. Senoo M, Mochida N, Wang L et al. Limited effect of chromatin remodeling on Dβ-to-Jβ recombination in CD4+CD8+ thymocyte: Implications for a new aspect in the regulation of TCR β gene recombination. Int Immunol 2001; 13:1405-1414.
43. Ferrier P, Krippl B, Blackwell TK et al. Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. EMBO J 1990; 9:117-125.
44. Eyquem S, Lagresle C, Fasseu M et al. Disruption of the lineage restriction of TCRβ gene rearrangements. Eur J Immunol 2002; 32:3256-3266.
45. Chattopadhyay S, Whitehurst CE, Chen J. A nuclear matrix attachment region upstream of the T-cell receptor β gene enhancer binds Cux/CDP and SATB1 and modulates enhancer-dependent reporter gene expression but not endogenous gene expression. J Biol Chem 1998; 273:29838-29846.
46. Whitehurst CE, Hu H, Ryu CJ et al. Normal TCRβ transcription and recombination in the absence of the Jβ2-Cβ2 intronic cis element. Mol Immunol 2001; 38:55-63.
47. Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the Dβ1 gene segment at the TCRβ locus by a germline promoter. Immunity 1999; 10:313-322.
48. Whitehurst CE, Schlissel MS, Chen J. Deletion of germline promoter PDβ1 from the TCRβ locus causes hypermethylation that impairs Dβ1 recombination by multiple mechanisms. Immunity 2000; 13:703-714.
49. Alvarez JD, Anderson SJ, Loh DY. V(D)J recombination and allelic exclusion of a TCR β-chain minilocus occurs in the absence of a functional promoter. J Immunol 1995; 155:1191-1202.
50. Halle JP, Haus-Seuffert P, Woltering C et al. A conserved tissue-specific structure at a human T-cell receptor β-chain core promoter. Mol Cell Biol 1997; 17:4220-4229.
51. Chen F, Rowen L, Hood L et al. Differential transcriptional regulation of individual TCR Vβ segments before gene rearrangement. J Immunol 2001; 166:1771-1780.
52. Ryu CJ, Haines BB, Lee HR et al. The T-cell receptor β variable gene promoter is required for efficient Vβ rearrangement but not allelic exclusion. Mol Cell Biol 2004; 24:7015-7023.
53. Sieh P, Chen J. Distinct control of the frequency and allelic exclusion of the Vβ gene rearrangement at the TCRβ locus. J Immunol 2001; 167:2121-2129.
54. Oestreich KJ, Cobb RM, Pierce S et al. Regulation of TCRβ gene assembly by a promoter/enhancer holocomplex. Immunity 2006; 24:381-391.
55. Hollenhorst PC, Shah AA, Hopkins C et al. Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. Genes Dev 2007; 21:1882-1894.
56. Kawazu M, Asai T, Ichikawa M et al. Functional domains of runx1 are differentially required for CD4 repression, TCRβ expression and CD4/8 double-negative to CD4/8 double-positive transition in thymocyte development. J Immunol 2005; 174:3526-3533.
57. Spicuglia S, Kumar S, Yeh JH et al. Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. Mol Cell 2002; 10:1479-1487.
58. Yang XO, Doty RT, Hicks JS et al. Regulation of T-cell receptor Dβ1 promoter by KLF5 through reiterated GC-rich motifs. Blood 2003; 101:4492-4499.
59. Rothenberg EV, Taghon T. Molecular genetics of T-cell development. Annu Rev Immunol 2005; 23:601-649.
60. Barndt R, Dai MF, Zhuang Y. A novel role for HEB downstream or parallel to the pre-TCR signaling pathway during αβ thymopoiesis. J Immunol 1999; 163:3331-3343.
61. Bender TP, Kremer CS, Kraus M et al. Critical functions for c-Myb at three checkpoints during thymocyte development. Nat Immunol 2004; 5:721-729.

62. Wakabayashi Y, Watanabe H, Inoue J et al. Bcl11b is required for differentiation and survival of αβ T-lymphocytes. Nat Immunol 2003; 4:533-539.
63. Kaul-Ghanekar R, Majumdar S, Jalota A et al. Abnormal V(D)J recombination of T-cell receptor β locus in SMAR1 transgenic mice. J Biol Chem 2005; 280:9450-9459.
64. Wolfer A, Wilson A, Nemir M et al. Inactivation of Notch1 impairs VDJβ rearrangement and allows pre-TCR-independent survival of early αβ lineage thymocytes. Immunity 2002; 16:869-879.
65. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. Cell 1985; 40:271-281.
66. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85:887-897.
67. Spicuglia S, Franchini DM, Ferrier P. Regulation of V(D)J recombination. Curr Opin Immunol 2006; 18:158-163.
68. Cobb RM, Oestreich KJ, Osipovich OA et al. Accessibility control of V(D)J recombination. Adv Immunol 2006; 91:45-109.
69. Mathieu N, Hempel WM, Spicuglia S et al. Chromatin remodeling by the T-cell receptor (TCR)-β gene enhancer during early T-cell development: implications for the control of TCR-β locus recombination. J Exp Med 2000; 192:625-636.
70. Tripathi R, Jackson A, Krangel MS. A change in the structure of Vβ chromatin associated with TCR β allelic exclusion. J Immunol 2002; 168:2316-2324.
71. Morshead KB, Ciccone DN, Taverna SD et al. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci USA 2003; 100:11577-11582.
72. Senoo M, Shinkai Y. Regulation of Vβ germline transcription in RAG-deficient mice by the CD3ε-mediated signals: implications of Vβ transcriptional regulation in TCR β allelic exclusion. Int Immunol 1998; 10:553-560.
73. Huang F, Cabaud O, Verthuy C et al αβ T-cell development is not affected by inversion of TCR β gene enhancer sequences: polar enhancement of gene expression regardless of enhancer orientation. Immunology 2003; 109:510-514.
74. Sikes ML, Meade A, Tripathi R et al. Regulation of V(D)J recombination: A dominant role for promoter positioning in gene segment accessibility. Proc Natl Acad Sci USA 2002; 99:12309-12314.
75. Krangel MS. T-cell development: better living through chromatin. Nat Immunol 2007; 8:687-694.
76. Osipovich O, Milley R, Meade A et al. Targeted inhibition of V(D)J recombination by a histone methyltransferase. Nat Immunol 2004; 5:309-316.
77. Osipovich O, Milley CR, Oestreich KJ et al. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrβ genes. Nat Immunol 2007; 8:809-816.
78. Baumann M, Mamais A, McBlane F et al. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. EMBO J 2003; 22:5197-5207.
79. Jackson A, Kondilis HD, Khor B et al. Regulation of T-cell receptor β allelic exclusion at a level beyond accessibility. Nat Immunol 2005; 6:189-197.
80. Jackson AM, Krangel MS. Allele-specific regulation of TCRβ variable gene segment chromatin structure. J Immunol 2005; 175:5186-5191.
81. Jia J, Kondo M, Zhuang Y. Germline transcription from T-cell receptor Vβ gene is uncoupled from allelic exclusion. EMBO J 2007; 26:2387-2399.
82. Corcoran AE. Immunoglobulin locus silencing and allelic exclusion. Semin Immunol 2005; 17:141-154.
83. Skok JA, Gisler R, Novatchkova M et al. Reversible contraction by looping of the Tcrα and Tcrβ loci in rearranging thymocytes. Nat Immunol 2007; 8:378-387.
84. Balomenos D, Balderas RS, Mulvany KP et al. Incomplete T-cell receptor Vβ allelic exclusion and dual Vβ-expressing cells. J Immunol 1995; 155:3308-3312.
85. Davodeau F, Peyrat MA, Romagne F et al. Dual T-cell receptor β chain expression on human T-lymphocytes. J Exp Med 1995; 181:1391-1398.
86. Padovan E, Giachino C, Cella M et al. Normal T-lymphocytes can express two different T-cell receptor β chains: implications for the mechanism of allelic exclusion. J Exp Med 1995; 181:1587-1591.
87. Mostoslavsky R, Alt FW, Rajewsky K. The lingering enigma of the allelic exclusion mechanism. Cell 2004; 118:539-544.
88. Gartner F, Alt FW, Monroe R et al. Immature thymocytes employ distinct signaling pathways for allelic exclusion versus differentiation and expansion. Immunity 1999; 10:537-546.
89. Khor B, Sleckman BP. Allelic exclusion at the TCRβ locus. Curr Opin Immunol 2002; 14:230-234.
90. Khor B, Sleckman BP. Intra- and inter-allelic ordering of T-cell receptor β chain gene assembly. Eur J Immunol 2005; 35:964-970.

91. Mostoslavsky R, Singh N, Tenzen T et al. Asynchronous replication and allelic exclusion in the immune system. Nature 2001; 414:221-225.
92. von Boehmer H, Aifantis I, Azogui O et al. Crucial function of the pre-T-cell receptor (TCR) in TCRβ selection, TCRβ allelic exclusion and αβ versus γδ lineage commitment. Immunol Rev 1998; 165:111-119.
93. Aifantis I, Buer J, von Boehmer H et al. Essential role of the preT-cell receptor in allelic exclusion of the T-cell receptor β locus. Immunity 1997; 7:601-607.
94. Ardouin L, Ismaili J, Malissen B et al. The CD3-γδε and CD3-ζ/η modules are each essential for allelic exclusion at the T-cell receptor β locus but are both dispensable for the initiation of V to (D)J recombination at the T-cell receptor-β, -γ and -δ loci. J Exp Med 1998; 187:105-116.
95. Aifantis I, Pivniouk VI, Gartner F et al. Allelic exclusion of the T-cell receptor β locus requires the SH2 domain-containing leukocyte protein (SLP)-76 adaptor protein. J Exp Med 1999; 190:1093-1102.
96. Uematsu Y, Ruser S, Dembic Z et al. In transgenic mice the introduced functional T-cell receptor β gene prevents expression of endogenous β genes. Cell 1988; 52:831-841.
97. Pircher H, Ohashi P, Miescher G et al. T-cell receptor (TcR) β chain transgenic mice: studies on allelic exclusion and on the TcR+ α/δ population. J Immunol 1990; 20:417-424.
98. Michie AM, Soh JW, Hawley RG et al. Allelic exclusion and differentiation by protein kinase C-mediated signals in immature thymocytes. Proc Natl Acad Sci USA 2001; 98:609-614.
99. Iritani BM, Alberola-Ila J, Forbush KA et al. Distinct signals mediate maturation and allelic exclusion in lymphocyte progenitors. Immunity 1999; 10:713-722.
100. Jackson AM, Krangel MS. A role for MAPK in feedback inhibition of tcrb recombination. J Immunol 2006; 176:6824-6830.
101. Eyquem S, Chemin K, Fasseu M et al. The Ets-1 transcription factor is required for complete pre-T-cell receptor function and allelic exclusion at the T-cell receptor β locus. Proc Natl Acad Sci USA 2004; 101:15712-15717.
102. Senoo M, Wang L, Suzuki D et al. Increase of TCR Vβ accessibility within Eβ regulatory region influences its recombination frequency but not allelic exclusion. J Immunol 2003; 171:829-835.
103. Suzuki D, Wang L, Senoo M et al. The positional effect of Eβ on Vβ genes of TCRβ chain in the ordered rearrangement and allelic exclusion. Int Immunol 2005; 17:553-1560.
104. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in Rag-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27:561-571.
105. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature advance online publication, 2007 (DOI 10.1038/nature06431).
106. Ramon-Maiques S, Kuo AJ, Carney D et al. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. Proc Natl Acad Sci USA 2007; 104:18993-18998.

Chapter 11

Molecular Pathways and Mechanisms Regulating the Recombination of Immunoglobulin Genes during B-Lymphocyte Development

Kristen Johnson, Karen L. Reddy and Harinder Singh*

Abstract

The hallmark of B-cell development is the ordered recombination of immunoglobulin (Ig) genes. Recently, considerable progress has been achieved in assembling gene regulatory networks comprised of signaling components and transcription factors that regulate B-cell development. In this chapter we synthesize experimental evidence to explain how such signaling pathways and transcription factors can orchestrate the ordered recombination of immunoglobulin (Ig) genes. Recombination of antigen-receptor loci is regulated both by the developmentally controlled expression of the *Rag1* and *Rag2* genes and the accessibility of particular loci and their gene segments to recombination. A new framework has emerged that invokes nuclear compartmentalization, large-scale chromatin dynamics and localized changes in chromatin structure in regulating the accessibility of Ig loci at specific stages of B-cell development. We review this emergent framework and discuss new experimental approaches that will be needed to explore the underlying molecular mechanisms.

Introduction

B-cell development is orchestrated by the coordinated action of signaling pathways and transcription factors that promote survival, proliferation and differentiation of hierarchically ordered progenitors (Fig. 1). B-lineage cells are directly derived from hematopoietic progenitors in the bone marrow termed early lymphoid progenitors (ELP) or common lymphoid progenitors (CLP) that also have the potential to differentiate into T, NK and to a lesser degree myeloid cells.[1,2] B-cell fate specification and commitment occur at the pro-B-cell stage. The rearrangement of immunoglobulin D-to-J_H gene segments is activated within the ELP/CLP pool of developmental intermediates and is completed at the pro-B-cell stage.[3,4] Since ELPs and CLPs can differentiate into alternate cell types, DJ_H rearrangements are not a defining molecular feature of B-lineage cells. However, the joining of V-to-DJ_H segments occurs exclusively within pro-B-cells and is a hallmark of commitment to the B-lineage. This step of rearrangement is highly regulated requiring multiple signaling and transcription factor inputs, thereby ensuring developmental timing and lineage specificity (Fig. 1). Productive rearrangement of an Ig heavy chain allele generates

*Corresponding Author: Harinder Singh—Howard Hughes Medical Institute, Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, IL, USA 60637, USA. Email: hsingh@uchicago.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

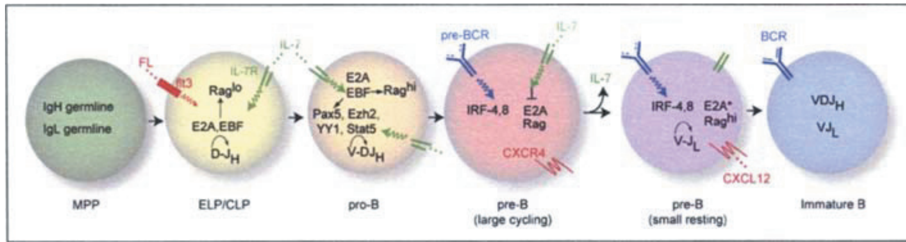


Figure 1. Overview of signaling pathways and transcription factors controlling B-cell development and V(D)J recombination. The initial and final configurations of the immunoglobulin loci are denoted in MPP and immature B-cells, respectively. Specific recombination events that occur in developmental intermediates are shown along with their regulators. In multipotent progenitors (MPP) that give rise to all hematopoietic lineages, the Ig loci are in their germline configuration (not recombined) and are transcriptionally inactive. Early lymphoid progenitors (ELP) or common lymphoid progenitors (CLP) express low levels of the transcription factors EBF and E2A that regulate the initial expression of the *Rag* genes and D-to-J_H recombination. IL-7 signaling has been implicated in regulating expression of EBF. B-cell fate specification is directed by the upregulation of EBF that induces Pax5 and results in the generation of pro-B-cells. IL-7 signaling via the transcriptional activator Stat-5 regulates distal V_H gene transcription and accessibility. In addition Pax-5, Ezh2 and YY1 are required for distal V-to-DJ_H recombination. After successful V-to -DJ_H recombination, the pre-BCR is displayed on the cell surface. Constitutive signaling through this receptor upregulates IRF4,8 expression while continued signaling through the IL-7 receptor initially inhibits V-to-Jκ recombination by blocking E2A accessibility to the intronic Igκ enhancer and repressing *Rag* gene expression. IRF4,8 directly bind to and activate the kappa 3′ enhancer. IRF4 also induces the chemokine receptor CXCR4 that is proposed to move pre-B-cells away from IL-7 producing stroma in the bone marrow resulting in loss of IL-7 signaling. Consequently, E2A access to the Igκ intronic enhancer (E2A*) is enabled, the *Rag* genes are highly expressed and efficient V-to-Jκ recombination ensues. Productive rearrangement of an Ig light chain allele results in assembly and expression of the B-cell receptor (BCR).

pre-B-cells in which the heavy-chain protein pairs with the surrogate light-chains, λ5 and Vpre-B, to form the pre-BCR.[5] Pre-B-cells undergo a self-limiting proliferative expansion mediated by the pre-BCR and the IL-7 receptor, thereby amplifying clones with successful IgH rearrangements. During this cycling pre-B-cell phase, the *Rag* genes are downregulated.[6,7] Upon cell cycle exit, *Rag* gene expression is re-induced and the cells activate rearrangement of their Ig light-chain loci. Both heavy and light chain rearrangements are subject to allelic exclusion, a process that ensures that only a single productively rearranged allele is generated and expressed thereby ensuring that a given B-cell expresses a unique antigen receptor. In this chapter we synthesize experimental evidence to understand how signaling pathways and transcription factors can orchestrate the developmentally ordered recombination of immunoglobulin (Ig) genes and also enforce allelic exclusion.

Recombination of antigen-receptor loci is regulated both by the developmentally controlled expression of the *Rag1* and *Rag2* genes that encode the V(D)J recombinase as well as by the accessibility of particular loci and their gene segments to recombination. Recently, a new framework has emerged that invokes nuclear compartmentalization and large-scale chromatin dynamics in addition to localized changes in chromatin structure in regulating the accessibility of Ig loci at specific stages of B-cell development. It has been established that the fundamental structural unit of chromatin, the nucleosome, can inhibit recombination when positioned directly over a recombination signal sequence (RSS).[8,9] Thus it is widely accepted that accessibility to recombination of individual gene segments must involve localized changes in nucleosome structure and positioning (Fig. 2).[10] Such changes are brought about by chromatin modifying complexes that are recruited to specific nucleosomal regions by transcription factors. Considerable progress has been achieved in elucidating distinct molecular mechanisms by which transcription factors and

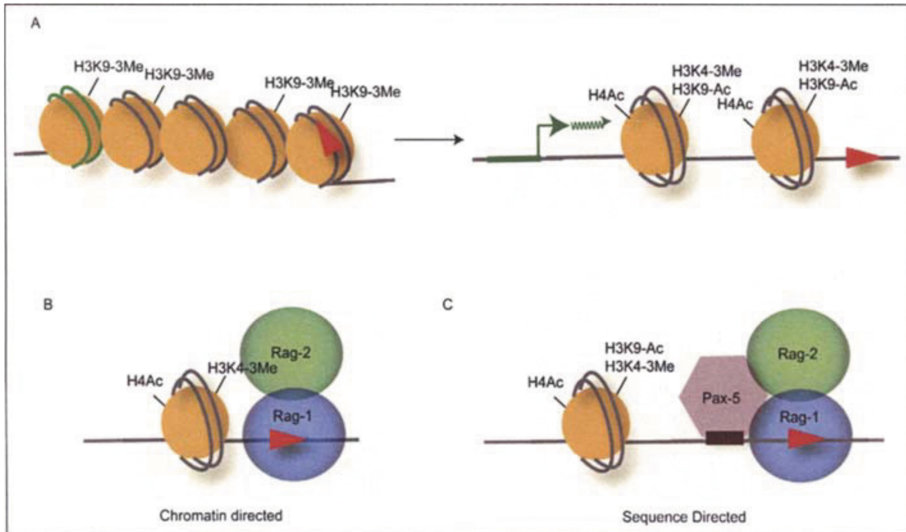


Figure 2. Molecular mechanisms regulating chromatin accessibility and recruitment of the RAG complex to recombination signal sequences (RSS). A) In the inactive state, Ig gene segments contain repressive H3K9 tri-methylation (H3K9-3Me) nucleosome modifications. This modification is proposed to inhibit recombinase accessibility to the RSS (red triangle). In this scenario, both the RSS and the V gene promoter (green DNA segment) are inaccessible. Activation of germline transcription (green arrow) is accompanied by nucleosome modifications that include histone H4 acetylation (H4-Ac), histone H3K9 acetylation (H3K9-Ac) and histone H3K4 tri-methylation (H3K4-3Me). This open chromatin structure is proposed to increase accessibility of RAG complex to the RSS. B) Chromatin directed RAG complex recruitment. This mechanism involves RAG-2 interaction with an RSS-proximal nucleosome containing H3K4-3Me. C) Sequence directed RAG complex recruitment. In this scenario the RAG complex is recruited by direct interactions with the transcription factor Pax-5 bound to a site adjacent to an RSS. Pax-5 interacts with both RAG-1 and RAG-2 proteins.

chromatin modifying complexes can locally regulate accessibility of antigen receptor gene segments to the V(D)J recombinase. Chromatin modifying complexes catalyze posttranslational modifications of histone tails that can serve to recruit nucleosome remodeling complexes that in turn alter the positioning of nucleosomes thereby directly changing the accessibility of recombination signal sequences (RSSs) to the V(D)J recombinase (Fig. 2A). In addition, the histone modifications can function as molecular scaffolds for more favorable binding of the RAG1,2 complex (Fig. 2B). Interestingly, a direct interaction between RAG2 and K4 trimethylated-histone H3 has recently been demonstrated.[11,12] This interaction appears to promote recombination in vivo. Transcription factors can also directly interact with the RAG1,2 proteins and recruit them to a nearby RSS (Fig. 2B). In this regard, the transcription factor Pax5 has been shown to bind to the RAG1 and RAG2 proteins and facilitate their recruitment to an RSS in vitro and promote recombination of a V_H gene substrate.[13] It is now recognized that developmentally regulated changes in nuclear compartmentalization of Ig loci also impact their accessibility to recombination (Fig. 3). Ig loci have been shown to associate in a regulated manner with two distinct repressive nuclear compartments, the inner nuclear membrane-nuclear lamina and pericentromeric heterochromatin, each of which appear to impair accessibility of these loci to recombination.[14-16] Given the fact that recombination of V with D or J gene segments often involve molecular synapses of RSSs separated by distances as large as 1-2Mb, an important issue is how such long-range recombination events are facilitated. Recent evidence suggests that Ig loci undergo compaction or contraction.[14,17,18] This phenomenon

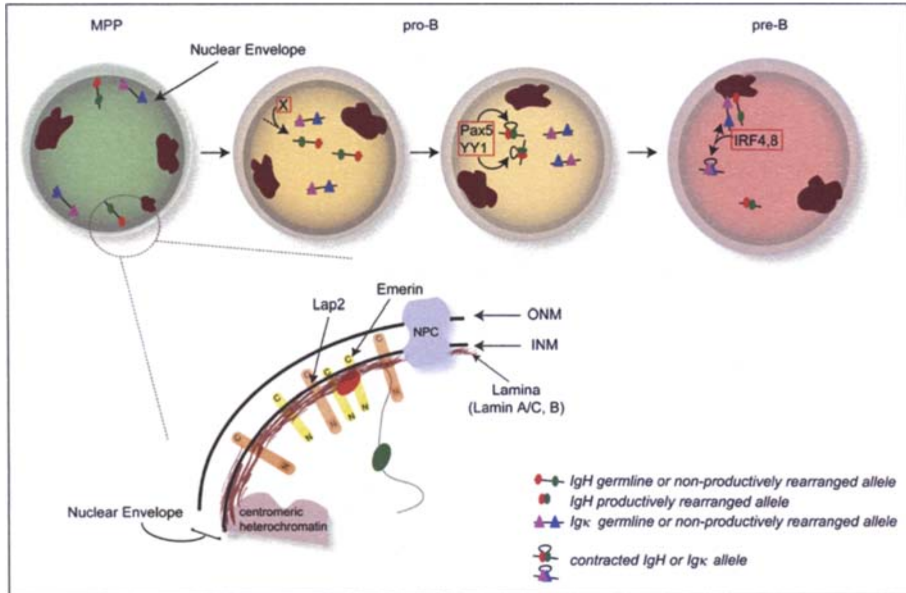


Figure 3. Regulation of Ig gene recombination via nuclear compartmentalization and DNA looping. Ig loci undergo developmentally regulated changes in nuclear compartmentalization that have been proposed to regulate recombination. These include associations with repressive compartments such as the inner nuclear membrane-lamina and pericentromeric heterochromatin. Ig loci also undergo large-scale changes in chromatin configuration, termed compaction or contraction that are thought to represent DNA loops which facilitate long range DNA recombination. The recombination status and nuclear disposition of the Ig heavy and kappa light chain alleles is depicted during various stages of B-cell development. The gray outline indicates the nuclear envelope that comprises the nuclear membrane and lamina. In multipotential progenitor cells or non-B-lineage cells (light green cell), both IgH and Igκ alleles are positioned at the nuclear periphery. At this stage, the IgH loci are decontracted and the distal V_H genes (red oval) are positioned closer to the periphery than the D_H, J_H or C_H regions (green oval). D-to-J_H recombination can occur in this state. Shown below are the proposed interactions of the IgH loci with the inner nuclear membrane (INM). ONM indicates outer nuclear membrane. The distal V_H genes, but not the C_H region, are in molecular contact with components of the INM-lamina including emerin and lamin B. As the cells transition to the pro-B-cell stage (yellow cells), the Ig loci are positioned away from the nuclear periphery by an unknown factor(s) X. In addition, at this stage, the IgH loci undergo compaction (contraction) thereby bringing the distal V_H gene segments in close proximity to the D_H-J_H region by looping out the intervening DNA. This contraction has been shown to be dependent on Pax-5, a positive regulator of V-to-DJ_H recombination and YY1 that binds to the Eμ enhancer in the C_H region. Finally, at the pre-B-cell stage (pink cell), one of the Igκ alleles undergoes contraction, while the other remains decontracted and is associated with pericentromeric heterochromatin. Moreover, the decontracted Igκ allele is often found to be associated with a decontracted IgH allele at the same pericentromeric heterochromatin cluster. Such association with pericentromeric heterochromatin is proposed to contribute to allelic exclusion. IRF4,8, are related transcription factors that are required for Igκ recombination. They regulate positioning of an Igκ allele away from pericentromeric heterochromatin and may promote contraction or DNA looping.

appears to reflect higher order chromosomal DNA loops that help to bring widely separated gene segments in closer proximity for DNA recombination. We review the new framework for analyzing Ig recombination accessibility at various levels, including nuclear compartmentalization, chromosome and chromatin structure.

B-Cell Fate Specification and the Joining of D-to-J_H Segments

D-to-J_H rearrangements are used to define hematopoietic progenitors in which lymphoid potential has been induced. Such recombination events are found at low levels in ELPs and CLPs.[1,2] Pro-B-cells invariably display DJ_H rearrangements at both IgH alleles. Genetic experiments have revealed a requirement for the cytokine receptors Flk2/Flt3 and IL-7R, as well as the transcription factors PU.1, Ikaros, E2A and EBF in the generation of pro-B-cells.[19] These signaling and transcription components have been assembled into contingent gene regulatory networks that initiate B-cell development. In this section we discuss the known functions of these regulators in the activation of DJ_H rearrangement and the molecular mechanisms underlying this earliest recombination event.

E2A and EBF are key regulators of B-cell fate specification and loss of either factor results in an early and profound block to B-cell development, in vivo, that appears to be at the stage involving initiation of D-to-J_H rearrangements.[20-22] Interestingly, ectopic expression of either transcription factor in conjunction with the RAG proteins in a nonlymphoid cell line is sufficient to activate D-to-J_H recombination.[23] Consistent with the view that either transcription factor can induce D-to-J_H recombination, it has been shown that neither E2A nor EBF are absolutely required for this rearrangement event during B-cell development. EBF can bypass the requirement for E2A in early B-cell development and induce D-J_H rearrangements in E2A mutant cells when expressed at sufficient levels.[24] Conversely, EBF mutant progenitors when propagated in the presence of FL and IL-7 express E2A and display D-to-J_H rearrangements.[25] The molecular mechanisms by which these factors are able to activate recombination of the DJ_H segments remain to be elucidated. An attractive possibility is that they bind to sites within and near the intronic IgH enhancer and locally remodel chromatin structure enabling accessibility of the nearby RSSs. In support of this possibility, the transactivating domain of E2A that interacts with chromatin modifying complexes is required to promote ectopic IgH recombination.[23]

Localized histone modifications have been implicated in the onset of D-to-J_H recombination on the basis of their selective appearance at the D-J_H locus as it is poised to undergo recombination (Fig. 4). In $CD19^+$ pro-B-cells isolated from *Rag* deficient mice, the D-J_H cluster is associated with hyperacetylated histones, suggesting a role for increased histone acetylation in creating a local region of accessibility that can be targeted by the recombinase machinery.[26] The concept of region-specific chromatin alterations as a means of effecting developmentally ordered changes in recombination accessibility has recently been strengthened by gene targeting studies that place a V_H segment in close proximity to the D_H elements. This resulted in a loss of ordered rearrangement for the targeted V_H gene segment.[27] Discrete chromatin domains within the IgH locus such as the one exemplified by the D-J_H cluster suggest the existence of boundary elements that may function to restrict recombination to gene segments within the domain.

Another process that is manifested at the DJ_H locus prior to recombination is antisense transcription.[28,29] Antisense intergenic transcription through this region is dependent on the Eμ enhancer, an element that has been shown to be required for D-to-J_H recombination.[29,30] Based on these results, it has been suggested that processive antisense transcription through the DJ_H region (60 kb) disrupts repressive chromatin structure thereby facilitating recombination.[29,31] Intriguingly, whereas the 5′- and 3′- D_H segments are associated with active H3-K9 acetylation marks the intervening D_H segments, that are infrequently recombined, appear to undergo active histone deacetylation.[28] It has been proposed that the intervening D_H genes are transcribed in both the antisense and sense orientations leading to the generation of low levels of dsRNA that promotes repeat induced epigenetic silencing. It remains to be determined if both of the fore-mentioned mechanisms are utilized in shaping the repertoire of D-to-J_H recombination events in pro-B-cells.

Despite the obvious requirement for D-to-J_H rearrangement during B-cell development, as noted above this step is less stringently regulated than V-to-DJ_H recombination. An intriguing explanation for this difference in regulated accessibility is suggested by the topology of the IgH locus in lymphoid progenitors. Specifically, the heavy-chain locus appears to be anchored at the nuclear lamina through

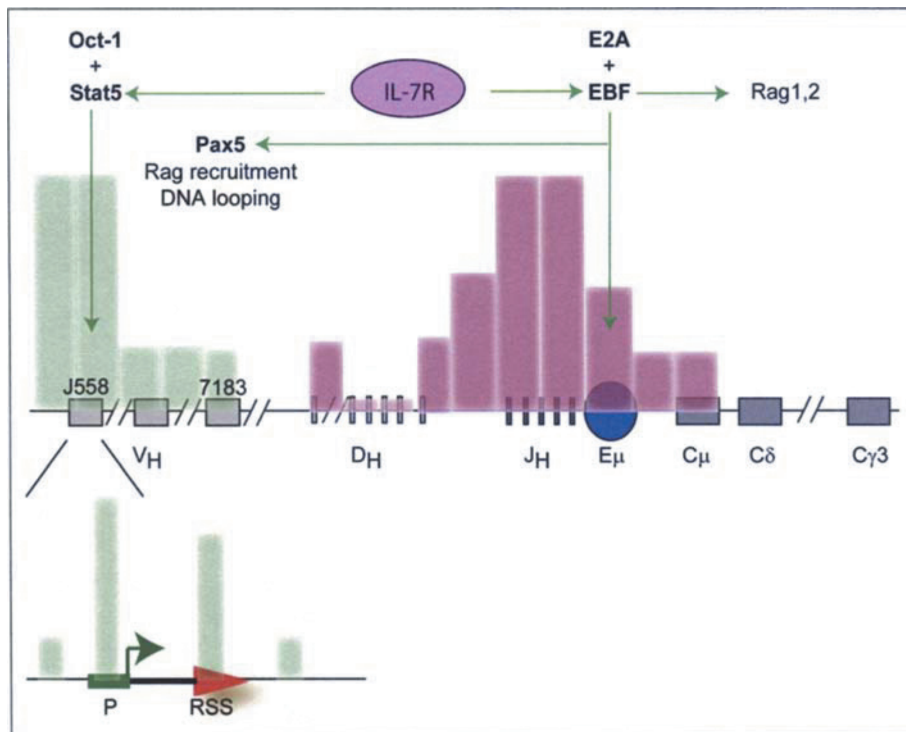


Figure 4. Signaling pathways and transcription factors regulating IgH recombination. The network depicts signaling pathways and transcriptional regulators that are required for heavy chain gene recombination in pro-B-cells. Gray boxes represent the indicated gene segments (not to scale). Early in pro-B-cell development, the C_H and D_H-J_H region adopts an open chromatin structure accompanied by localized histone acetylation (pink bars). Recombination of the D-to-J_H segments is proposed to influence the local chromatin structure and accessibility of proximal V_H genes. After D-to-J_H recombination, the V_H domain becomes differentially acetylated, with the more distal gene segments displaying relatively higher levels histone H4 acetylation. IL-7 signaling via Stat-5 and its interaction with Oct-1 regulates acetylation of the distal V_H genes. Pax-5 is dispensable for the acetylation and transcription of the distal V_H gene segments, but is required for their compaction and recombination. Importantly, H4 acetylation is localized to nucleosomes positioned near V_H gene promoters (dark green) and RSS (red triangle) but not in the intergenic regions.

the distal V_H domain.[14,16,17,32] As the locus is in an extended conformation at this stage of development, the V_H gene segments are more closely associated with a repressive compartment than the DJ_H cluster (Fig. 3). Consequently, the DJ_H genes segments may be more accessible to the V(D)J recombinase than the V_H segments (see below). Though the role of nuclear topology of the IgH locus in differentially regulating its accessibility remains to be rigorously established, these analyses encompassing transcription factors, chromatin structure, antisense transcription and nuclear organization highlight both local and global mechanisms that likely regulate recombination.

B-Cell Fate Commitment and V-to-DJ_H Rearrangement

The transcription factor EBF induces B-cell fate specification and also initiates B-cell fate commitment by restricting alternate myeloid lineage options.[25,33] IL-7R signaling is required for the developmental induction of the EBF gene in lymphoid progenitors (Fig. 1).[34] EBF in turn induces

the expression of Pax-5, a transcription factor that is required for B-cell fate commitment.[19,35] EBF and Pax5 are essential for the generation of committed pro-B-cells in which V-to-DJ_H rearrangements are activated. It is important to note that these B-lineage specific recombination events at the IgH locus not only coincide with B-cell fate commitment but they are directly regulated by signaling pathways (IL-7R) and transcription factors (EBF and Pax5) that are needed to establish the committed state.

The V_H domain of the IgH locus spans approximately 2 Mb and includes approximately 150 functional gene segments, each of which has its own promoter and RSS element.[36,37] The gene segments are grouped into families based on sequence homology. Distinct regulatory pathways and mechanisms are involved in controlling the recombination accessibility of D_H proximal versus D_H distal V_H genes. During B-cell development in the fetal liver, the D_H proximal V_H gene segments are preferentially recombined.[38-41] This selective rearrangement of V_H genes is considered to be a consequence of their closer proximity to the DJ_H segments. Intriguingly, it has been demonstrated that the D_H proximal portion of the V_H domain becomes associated with hyperacetylated histones following successful DJ_H rearrangement. These results suggest an attractive mechanism involving the limited spreading of activating histone marks to account for the preferential rearrangement of proximal V_H genes seen early in development (Fig. 4).[26] They also provide a means for sequentially ordering the recombination of D-to-J_H segments and proximal V-to-DJ_H gene segments within the IgH locus. Recently, it has been shown that the transcription factor EBF is required for V-to-DJ_H recombination.[25] It will be important to determine if EBF targets proximal V_H genes and regulates their chromatin structure thereby promoting their recombination.

Considerable progress has been achieved in analyzing the recombination of the distal V_H genes. The IL-7 signaling pathway specifically regulates recombination of distal V_H gene segments.[42,43] Locally restricted histone acetylation associated with the individual distal V_H gene segments is dependent on IL-7 signaling.[26,44] The transcription factor STAT5, a downstream signaling component of the IL-7 signaling pathway has been shown to be required for efficient distal V_H gene rearrangement, thus establishing a molecular link between the IL-7 signaling pathway and IgH recombination.[43] Stat5 is recruited to V_H gene promoters via the transcription factor Oct-1 that binds to the conserved octamer element. Stat5 functions as a co-activator, stimulating germline transcription, histone acetylation and recombination of the distal V_H gene segments.[43] It should be noted that Stat5 is not required for the repositioning of IgH alleles away from the nuclear lamina or for their compaction, both of these higher-order steps are also implicated in regulating distal V_H gene recombination (see below). Therefore, IL-7 signaling via the transcription factor Stat5 appears to function specifically in regulating localized changes in accessibility of distal V_H gene segments through histone modifications and possibly nucleosome remodeling.

A second key regulator of distal V_H gene recombination is the transcription factor Pax5 (BSAP). Importantly, Pax5 regulates the recombination of distal V_H gene segments via a molecular mechanism that is distinct from the one detailed above for Stat5.[43] In the absence of Pax5, B-cell development is arrested at the pro-B-cell stage.[45] In contrast to the block seen in Stat5 deficient cells, the distal V_H gene segments are associated with highly acetylated histones and undergo normal levels of germline transcription in the absence of Pax5.[46] Although, the IgH alleles are centrally positioned in the nuclei of Pax5 mutant pro-B-cells, they do not undergo compaction, also termed contraction (Fig. 3).[17] Importantly, restoration of Pax5 expression in Pax5, Rag2 mutant cells induces contraction of IgH alleles. These results demonstrate that Pax5 can promote IgH locus contraction thereby increasing the spatial proximity of distal V_H gene segments and the DJ_H cluster, in the absence of recombinase activity. Intriguingly, Pax5 is also implicated in the loss of H3-K9 methylation in pro-B-cells by promoting exchange with the histone variant H3.3 (see below).[47] We note that Pax5 has been shown to bind to multiple V_H gene segments in B-lineage cells.[13] Moreover, as noted above, Pax5 interacts with the RAG1,2 proteins.[13] Thus Pax5 appears to regulate distal V_H gene recombination via multiple mechanisms that include removal of repressive histone modifications, promoting locus contraction and directly recruiting RAG1,2 complexes (Figs. 2,3).

IgH locus compaction or contraction is a manifestation of higher-order chromosomal DNA loops that juxtapose distal V_H gene segments with the DJ_H cluster and has been proposed to promote distal V_H gene recombination.[17,18] Interestingly, ectopic expression of Pax5 in T-lineage cells induces V-to-DJ_H recombination but paradoxically results in a similar IgH recombination phenotype as seen in Pax5 mutant pro-B-cells. Pax5 mis-expressing T-lineage cells rearrange proximal but not distal V_H gene segments despite the fact that these gene segments are highly transcribed and the IgH loci are centrally located.[17,48] It should be noted that Pax5 expressing T-lineage cells can also activate the EBF gene and the latter factor may account for their ability to undergo proximal V-to-DJ_H recombination. These results have led to the suggestion that Pax5 acts in conjunction with another B-cell specific factor to induce IgH locus contraction and facilitate distal V-to-DJ_H rearrangement.

The zinc finger transcription factor YY1 has also recently been shown to be involved in IgH locus contraction.[49] B-lineage specific deletion of the YY1 gene results in a block to development at the pro-B-cell stage and impaired V-to-DJ_H recombination that is most severely manifested for distal V_H gene segments.[49] YY1 binds to the heavy-chain intronic enhancer and has been proposed to play a direct role in locus contraction i.e., DNA looping by promoting enhancer-promoter interactions. Importantly, Pax5 expression is not altered in YY1 mutant pro-B-cells. Thus Pax5 and YY1 are independently required for IgH locus contraction. Locus contraction is likely to facilitate molecular synapsis of two widely separated and compatible RSSs by the RAG protein complexes. Once widely separated gene segments are brought into proximity with one another, RAG proteins can then achieve molecular synapsis. RAG proteins have been inferred to preferentially associate with RSS elements containing a 12 bp spacer rather than with an RSS containing a 23 bp spacer in vivo.[50] These data support the "capture" model, which posits that oligomeric RAG complexes initially bind to an RSS containing gene segment and then capture the complementary RSS containing segment leading to molecular synapsis and the initiation of recombination via DNA cleavage.

In addition to the fore-mentioned transcription factors, the polycomb group protein Ezh2 is also required for rearrangement of the distal V_H gene segments.[51] Strikingly, the molecular phenotype of Ezh2 mutant pro-B-cells is very similar to that of Pax5 mutant cells in that distal V_H gene segments are highly transcribed and associated with hyperacetylated histones despite the block to their recombination.[46,51] Ezh2 is a histone methyltransferase, that methylates histone H3 at K27. Loss of Ezh2 results in reduced H3 methylation at distal V_H gene segments.[51] It remains to be determined if Ezh2 as is the case for Pax5 and YY1 is required for IgH locus contraction.

There is an additional developmentally regulated chromatin modification, H3-K9 methylation, which appears to regulate heavy-chain recombination.[47] H3-K9 methylation is associated with the V_H locus in non-B-lineage cells but is removed in pro-B-cells. H3-K9 methylation has been demonstrated to severely inhibit recombination upon its targeting to an engineered RSS containing substrate in a B-cell line[52]. Interestingly, as is case for the DJ_H region, antisense transcription also occurs at the V_H locus and may be involved in promoting exchange of repressive histones.[53]

The above studies enable us to propose a model for the developmental control of immunoglobulin heavy chain recombination in developing B-cells. In non-B-cells and hematopoietic progenitors the germline heavy-chain alleles are associated with the inner nuclear membrane-nuclear lamina compartment and assembled in a repressive chromatin structure involving H3-K9 methylation (Fig. 2). As the V_H gene segments are more closely interacting with the INM-lamina than the DJ_H cluster, the latter segments may undergo recombination while positioned at the nuclear periphery (Fig. 3). B-cell fate specification requires the transcription factors E2A and EBF. In lymphoid progenitors, these two factors appear to regulate the initial low-level expression of the *Rag* genes as well as the accessibility of the DJ_H cluster to recombination, likely by promoting antisense transcription and chromatin modifications. Increased expression of EBF as a consequence of IL-7R signaling promotes B-cell fate specification and the generation of committed pro-B-cells via the induction of Pax5. During B-cell fate specification the Ig heavy-chain alleles are repositioned away from the nuclear lamina through as yet unidentified factors and mechanisms.[14] EBF is an attractive candidate regulator for inducing repositioning of IgH alleles. In *EBF-/-* lymphoid progenitors

the D-J recombined IgH alleles are positioned at the nuclear lamina (I. DeMarco and H. Singh, unpublished results). In pro-B-cells, IL-7 signaling via Stat5, induces localized chromatin alterations in distal V_H gene segments and activates their transcription.[43] Pax5 along with YYI promotes IgH locus contraction facilitating the recombination of distal V_H gene segments.[17,49] A recent study compared the distance distributions of FISH signals from multiple small probes (10kb) that hybridize to the IgH locus and used computer modeling and triangulation to demonstrate that the locus is organized into compartments containing clusters of loops separated by linkers.[54] Importantly, in pro-B-cells, the entire 2Mbp region containing the V_H genes appears to be juxtaposed to the D_H elements, thus facilitating long-range genomic interactions.[54] It will be important to determine how the transcription factors Pax5 and YY1 that appear to impact distinct domains of the IgH locus contribute to its structural reconfiguration in pro-B-cells. The molecular functions of the transcription factors EBF, Pax5 and YY1 in regulating antisense V_H transcripts remain to be explored.[53] We note that proximal V_H gene rearrangement requires EBF but not Pax5 or YY1.[25,46,49] Thus regulated chromatin alterations, interactions with the INM-lamina compartment that are domain specific and locus reconfiguration accompanied by compaction appear to promote the accessibility of the large repertoire of V_H gene segments to recombination in developing B-cells.

The Pre-B-Cell Checkpoint and the Induction of Light-Chain Recombination

Following productive heavy-chain rearrangement B-cells progress through a critical developmental checkpoint (Fig. 1). This process consists of a self-limiting clonal expansion culminating in cell cycle exit and initiation of light-chain rearrangement. Successful light-chain rearrangement results in the generation of immature IgM+ B-cells. The *Rag* genes are down regulated during the cycling pre-B-cell phase and re-induced upon cell cycle exit. Signaling through the pre-BCR and the IL-7R regulates the pre-B to B-cell transition. IL-7 signaling is active during the large cycling pre-B-cells stage. However, the pre-BCR reduces the dependence of pre-B-cells on IL-7 and this is also correlated with a change in the anatomic distribution of pro-B and pre-B-cells in the bone marrow. The former are associated with IL-7 expressing stroma whereas the latter are positioned away from such stromal cells.[55,56] One of two light-chain loci, Igκ or Igλ, undergo productive rearrangement at the pre-B-cell stage. Their genomic structures are depicted in Figure 5. In mice, the Igκ locus is more frequently rearranged, at a ratio of 20:1 and consequently recombination of this locus has been more intensively studied.[57]

Signaling through both the pre-BCR and the IL-7R drives the limited clonal expansion of pre-B-cells. *Rag* gene expression is down regulated during this phase (Fig. 1).[7,22] Therefore, the proliferative burst separating the IgH and IgL recombination events during B-cell development provides pre-B-cells with the opportunity to pause recombination in the absence of an active recombinase and redirect chromatin accessibility from the heavy-chain locus to the light-chain loci. Until recently, it was considered that cell-cycle exit, may be sufficient to initiate Ig light-chain recombination. However, a combination of loss-of-function and gain-of-function experiments involving key cell cycle regulators, have demonstrated that exit of pre-B-cells from the cell-cycle is not a sufficient condition for the activation of recombination.[58] Instead, acquired pre-BCR signaling followed by attenuated IL-7R signaling results in alteration of chromatin accessibility of Ig light chain loci and cell cycle exit. Attenuation of IL-7R signaling also contributes to the optimal expression of the *Rag* genes and high recombinase activity.

An area of intense investigation concerning the regulation of Ig light-chain recombination has involved the analysis of transcription factors that bind to and activate transcriptional enhancers within the Igκ locus. The simplest explanation for the restriction of light-chain recombination to the pre-B-cell stage would be developmentally appropriate expression of Igκ-specific transcription factors. As detailed below, the molecular mechanism is not quite that simple. Nevertheless, recent insight suggests an exquisitely regulated process that integrates the developmental signaling programs found in pre-B-cells to the activities of key transcription factors ultimately leading to stage-specific Igκ recombination.

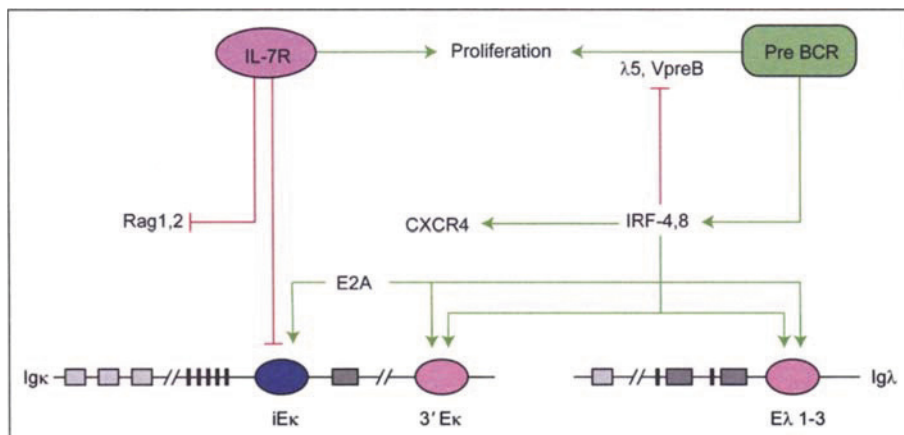


Figure 5. Signaling pathways and transcription factors regulating Igκ and Igλ recombination. The network depicts signaling pathways and transcriptional regulators that are required for light-chain recombination at the pre-B-cell stage. Arrows represent positive regulation and barred lines represent repression. As indicated, IRF-4 plays a central role in inducing light-chain recombination downstream of the pre-BCR by directly engaging the 3'Eκ and λ light-chain enhancers. IRF-4,8 are also suggested to induce cell cycle arrest and modulate IL-7 signaling thereby resulting in robust induction of *Rag* gene expression and E2A binding to iEκ. IRF4,8 induce chemokine receptors that are proposed to induce migration of pre-B-cells away from IL-7 producing stroma, leading to attenuation of IL-7 signaling and activation of the Igκ enhancer (through E2A binding) as well as increased *Rag* gene expression.

Genetic analyses have demonstrated that the transcription factors, E2A, Pax5 and the related family members IRF-4 and IRF-8 are required for light-chain recombination (Fig. 5).[59-61] These factors have known binding sites within the Ig kappa enhancers and in vivo DNA footprinting analysis has shown that their sites are occupied in primary pre-B-cells.[62] Interestingly, footprinting analysis comparing the binding of these key transcription factors during the transition from pro-B to pre-B-cells demonstrates no change with the exception of the composite site for PU-1/IRF-4.[62] Interestingly, IRF-4 expression increases at the pre-B-cell stage.[58,63,64] Loss of IRF-4 along with IRF-8, results in a complete block to B-cell development at the large cycling pre-B-cell stage with a failure to undergo Igκ or Igλ recombination.[60] A detailed analysis of the molecular mechanisms by which IRF-4 and IRF-8 activate recombination of Ig light-chain loci by is provided below. Unlike IRF-4 and IRF-8, the transcription factors Pax5 and E2A also function earlier in B-cell development at the pro-B-cell stage, where they are required for Ig heavy-chain recombination. Using Pax5 deficient or E2A deficient pre-B-cells, it has been shown that both factors additionally regulate Igκ germline transcription and recombination.[59,61]

Signaling by the pre-BCR has been widely considered to activate light-chain recombination. Expression of a transgene encoding the Igμ heavy-chain protein increases Igκ locus accessibility in *Rag* deficient pro-B-cells.[65-67] Additionally, the enforced expression of activated Ras, a downstream signaling component of the pre-BCR, promotes Ig light-chain recombination in the absence of an Ig heavy-chain.[68] Conversely, loss of signaling molecules including BLNK, Btk and PCLγ, that lie downstream of the pre-BCR, results in fewer cells that have rearranged their Ig kappa loci.[69,70] As noted above, the transcription factor IRF-4 is induced by pre-BCR signaling and Ig light-chain recombination is blocked in *Irf-4,8*$^{-/-}$ pre-B-cells despite the high expression of the pre-BCR.[60] Restoring either IRF-4 or IRF-8 expression rescues developmental progression and activates Ig light-chain rearrangement.[58,71] IRF-4 promotes histone acetylation at critical enhancers within Igκ and Igλ loci and induces their germline transcription (Fig. 5).[58] Intriguingly, IRF-4 also counteracts association of an Igκ allele with pericentromeric heterochromatin, an interaction that has been

proposed to inhibit recombination (Fig. 3).[58] These data delineate a molecular pathway by which pre-BCR signaling regulates both Igκ and Igλ recombination and also provide insights into the underlying molecular mechanisms.

Several studies have implicated IL-7 signaling in the negative regulation of Ig light-chain recombination. Withdrawal of IL-7 in pro-B-cell cultures appears to induce Ig light-chain recombination.[72] However, Ig light-chain recombination can occur in the presence of high concentrations of IL-7 and it has been argued that IL-7 withdrawal merely selects for cells that have undergone productive light-chain recombination.[73] Until recently the precise role of IL-7 signaling in regulating Ig light-chain recombination had remained unclear.[55,73] Using *Irf-4,8*$^{-/-}$ pre-B-cells, it has been demonstrated that IL-7 signaling can regulate Ig light chain recombination independently of pre-BCR signaling and IRF-4. Attenuating IL-7 signaling in *Irf-4,8*$^{-/-}$ pre-B-cells activates Igκ but not Igλ recombination.[58] Recombination is accompanied by the induction of Igκ germline transcripts and substantial upregulation of *Rag* transcripts. Intriguingly, binding of E2A to the intronic Igκ enhancer and localized histone acetylation increases within 24 hours of attenuated IL-7 signaling. Thus IL-7 signaling modulates Igκ rearrangement in pre-B-cells by controlling the activity of the intronic Igκ enhancer as well as optimal expression of the *Rag* genes. As IL-7 signaling is active in pro-B-cells it would inhibit Ig light-chain recombination at this stage. As noted above, IL-7 signaling promotes Ig heavy chain rearrangement in pro-B-cells and this pathway is dependent on Stat5. It remains to be determined if inhibition of Igκ rearrangement by IL-7 signaling is also dependent on Stat5 and if so what is the nature of the molecular mechanism by which Stat5 regulates accessibility of E2A at the intronic Igκ enhancer.

Despite the fact that pre-BCR and IL-7 signaling pathways can function independently of one another in promoting Igκ recombination, it is highly likely that their activities are coordinated during B-cell development. Consistent with this view, the two pathways function synergistically to induce Igκ recombination and the generation of IgM expressing B-cells.[58] The molecular basis of synergy in promoting Igκ recombination appears to be manifested at two steps. Firstly, each pathway targets a distinct Igκ enhancer and synergy is likely a consequence of simultaneously activating both enhancers. Secondly, IRF-4 preferentially induces Igκ germline transcription whereas attenuation of IL-7 signaling more highly induces *Rag* gene expression thereby optimizing changes in accessibility with expression of the recombinase.

An intriguing model has been proposed for the regulation of Ig light chain recombination via the coordination of pre-BCR and IL-7 signaling pathways in vivo. Genome-wide expression analysis using *Irf-4,8*$^{-/-}$ pre-B-cells revealed a number of genes involved in cell migration and adhesion that are regulated by IRF-4.[58] Of particular interest was the gene encoding CXCR4, a chemokine receptor that promotes migration in response to CXCL12. IRF-4 dependent upregulation of CXCR4 was shown to result in a change in the chemotactic properties of pre-B-cells. Since CXCL12 expressing stromal cells are spatially separated from IL-7 expressing stromal cells in the bone marrow, it has been proposed that IRF-4 regulated chemotaxis towards CXCL12 expressing stomal cells results in repositioning of pre-B-cells away from the IL-7 expressing stroma.[56,58] This movement would result in attenuation of IL-7 signaling and promote the synergistic induction of Ig light-chain recombination by the two molecular pathways detailed above.

Allelic Exclusion

Allelic exclusion of both IgH and IgL loci ensures the generation of B-cells that express a single type of antigen receptor. For each locus productive rearrangement of one allele culminates in feed back inhibition of further rearrangement of the other allele. We will initially discuss the molecular mechanisms that have been suggested to regulate allelic exclusion of the Igκ locus, as it has been more intensively studied. Allelic exclusion at the Igκ locus is initiated by a single allele being chosen to undergo recombination at the pre-B-cell stage. Two fundamentally different mechanisms, stochastic versus directed, have been proposed to explain this phenomenon. The stochastic mechanism invokes limiting amounts of either a transcription factor(s) that regulates locus accessibility or limiting expression of the recombination machinery. Either condition is

proposed to lead to inefficient recombination thereby decreasing the probability that both alleles undergo recombination simultaneously.[74,75] Data in support of this mechanism has been obtained by monitoring GFP expression in a knock-in mouse that expresses a GFP cDNA from an unrearranged κ allele.[76] Only a small percentage of pre-B-cells were seen to express GFP and such Igκ germline transcription was monoallelic. This data has been interpreted to suggest that a limiting transcription factor that activates Igκ germline transcription in pre-B-cells also restricts recombination to the small fraction of activated alleles. An alternative explanation for allelic exclusion of the Igκ locus proposes a series of directed epigenetic changes that occur differentially on individual kappa alleles. In agreement with this hypothesis, tight correlations have been found between monoallelic DNA demethylation of Igκ alleles and their replication timing.[76-79] More detailed analyses have revealed that at the pre-B-cell stage the early replicating Igκ allele is assembled into an active chromatin structure and preferentially undergoes DNA demethylation thereby increasing its accessibility to recombination.[80] In contrast, the late replicating allele is assembled into inactive chromatin comprising hypoacetylated histones and methylated H3-K9. This allele is also associated with pericentromeric heterochromatin and suggested to be a poorer substrate for recombination. Intriguingly, a cis-element, termed Sis, has been discovered in the V-Jκ intervening sequence and this element targets an Igκ transgene to pericentromeric heterochromatin.[81] Using yeast artificial chromosome-based single copy transgenic mice the Sis element was shown to negatively regulate Igκ recombination.[81] Moreover, this element was shown to interact with the zinc finger protein Ikaros, a transcription factor that has been shown to be associated with transcriptionally inactive genes, including κ allele that are associated with pericentromeric heterochromatin.[15,80] These data have led to the suggestion that Ikaros-Sis complexes actively participate in the process of allelic exclusion by promoting silencing of a single Igκ allele via interaction with pericentromeric heterochromatin. These distinct sets of observations concerning monoallelic activation of the Igκ locus have utilized different methodologies and cannot be easily reconciled. It is possible that a directed mechanism is used to distinguish the two alleles and a limiting transcription factor further restricts the activation of the more accessible allele to a small percentage of pre-B-cells.

Allelic exclusion at the heavy-chain locus involves feedback inhibition by the product of the productively rearranged allele (assembled into the pre-BCR) and attenuation of IL-7 receptor signaling.[5,82] It has been shown that the nonproductively rearranged heavy-chain allele is recruited to pericentromeric heterochromatin and undergoes locus decontraction following successful rearrangement of the other allele.[83] Recently, an intriguing mechanism involving specific inter-chromosomal interactions between the heavy-chain and light-chain loci has been proposed to link allelic exclusion at both loci.[84] Using 3D FISH, IgH and Igκ alleles were found to colocalize with pericentromeric heterochromatin in pre-B-cells. This inter-chromosomal interaction was dependent on the Ig 3'κ enhancer. Deletion of this cis-regulatory element resulted in not only loss of the association between IgH and Igκ alleles but prevented IgH locus decontraction. This was suggested to promote continued accessibility of the Ig heavy-chain locus to recombination in pre-B-cells and a breakdown of allelic exclusion.

Perspectives

The analysis of transcription factors and signaling pathways that regulate immunoglobulin gene recombination during B-lymphocyte development has resulted in considerable progress. A plausible developmental scheme can now be formulated for the ordered recombination of Ig heavy and light chain loci. The transcription factors not only appear to regulate Ig locus accessibility via localized changes in chromatin structure but also likely modulate recombination by altering nuclear compartmentalization of Ig alleles and their large-scale chromatin dynamics. Future research should uncover novel molecular components that mediate the interactions of Ig loci with the INM-lamina compartment or pericentromeric heterochromatin and test if they regulate recombination. Furthermore, the molecular mechanisms underlying large-scale DNA loops at Ig loci remain to be elucidated. Formation of these intrachromosomal loops is likely to be required for long-range V(D)J recombination and the generation of a diverse repertoire of antigen receptors.

References

1. Igarashi H, Gregory SC, Yokota T et al. Transcription from the RAG1 locus marks the earliest lymphocyte progenitors in bone marrow. Immunity 2002; 17(2):117-130.
2. Allman D, Miller JP. Common lymphoid progenitors, early B-lineage precursors and IL-7: characterizing the trophic and instructive signals underlying early B-cell development. Immunol Res 2003; 27 (2-3):131-140.
3. Li YS, Hayakawa K, Hardy RR. The regulated expression of B-lineage associated genes during B-cell differentiation in bone marrow and fetal liver. J Exp Med 1993; 178(3):951-960.
4. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell 2002; 109 Suppl:S45-55.
5. Geier JK, Schlissel MS. Pre-BCR signals and the control of Ig gene rearrangements. Semin Immunol 2006; 18(1):31-39.
6. Lin H, Grosschedl R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature 1995; 376(6537):263-267.
7. Li Z, Dordai DI, Lee J et al. A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. Immunity 1996; 5(6):575-589.
8. Golding A, Chandler S, Ballestar E et al. Nucleosome structure completely inhibits in vitro cleavage by the V(D)J recombinase. EMBO J 1999; 18(13):3712-3723.
9. Kwon J, Imbalzano AN, Matthews A et al. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. Mol Cell 1998; 2(6):829-839.
10. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
11. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450(7172):1106-1110.
12. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27(4):561-571.
13. Zhang Z, Espinoza CR, Yu Z et al. Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V(H)-to-DJ(H) rearrangement of immunoglobulin genes. Nat Immunol 2006; 7(6):616-624.
14. Kosak ST, Skok JA, Medina KL et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 2002; 296(5565):158-162.
15. Skok JA, Brown KE, Azuara V et al. Nonequivalent nuclear location of immunoglobulin alleles in B-lymphocytes. Nat Immunol 2001; 2(9):848.
16. Reddy KL, Zullo JM, Bertolino E et al. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature 2008; 452(7184):243-247.
17. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18(4):411-422.
18. Sayegh C, Jhunjhunwala S, Riblet R et al. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B-cells. Genes Dev 2005; 19(3):322-327.
19. Medina KL, Singh H. Gene regulatory networks orchestrating B-cell fate specification, commitment and differentiation. Curr Top Microbiol Immunol 2005; 290:1-14.
20. Bain G, Maandag EC, Izon DJ et al. E2A proteins are required for proper B-cell development and initiation of immunoglobulin gene rearrangements. Cell 1994; 79(5):885-892.
21. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B-cell formation. Cell 1994; 79(5):875-884.
22. Lin WC, Desiderio S. V(D)J recombination and the cell cycle. Immunol Today 1995; 16(6):279-289.
23. Romanow WJ, Langerak AW, Goebel P et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Mol Cell 2000; 5(2):343-353.
24. Seet CS, Brumbaugh RL, Kee BL. Early B-cell factor promotes B-lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. J Exp Med 2004; 199(12):1689-1700.
25. Pongubala JM, Northrup DL, Lancki DW et al. Transcription factor EBF restricts alternative lineage options and promotes B-cell fate commitment independently of Pax5. Nat Immunol 2008; 9(2):203-215.
26. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J 2001; 20(22):6394-6403.
27. Bates JG, Cado D, Nolla H et al. Chromosomal position of a VH gene segment determines its activation and inactivation as a substrate for V(D)J recombination. J Exp Med 2007; 204(13):3247-3256.
28. Chakraborty T, Chowdhury D, Keyes A et al. Repeat organization and epigenetic regulation of the DH-Cmu domain of the immunoglobulin heavy-chain gene locus. Mol Cell 2007; 27(5):842-850.
29. Bolland DJ, Wood AL, Afshar R et al. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. Mol Cell Biol 2007; 27(15):5523-5533.

30. Afshar R, Pierce S, Bolland DJ et al. Regulation of IgH gene assembly: role of the intronic enhancer and 5′DQ52 region in targeting DHJH recombination. J Immunol 2006; 176(4):2439-2447.
31. Chakalova L, Debrand E, Mitchell JA et al. Replication and transcription: shaping the landscape of the genome. Nat Rev Genet 2005; 6(9):669-677.
32. Yang Q, Riblet R, Schildkraut CL. Sites that direct nuclear compartmentalization are near the 5′ end of the mouse immunoglobulin heavy-chain locus. Mol Cell Biol 2005; 25(14):6021-6030.
33. Kikuchi K, Lai AY, Hsu CL et al. IL-7 receptor signaling is necessary for stage transition in adult B-cell development through up-regulation of EBF. J Exp Med 2005; 201(8):1197-1203.
34. Dias S, Silva H Jr, Cumano A et al. Interleukin-7 is necessary to maintain the B-cell potential in common lymphoid progenitors. J Exp Med 2005; 201(6):971-979.
35. Busslinger M. Transcriptional control of early B-cell development. Annu Rev Immunol 2004; 22:55-79.
36. Chevillard C, Ozaki J, Herring CD et al. A three-megabase yeast artificial chromosome contig spanning the C57BL mouse Igh locus. J Immunol 2002; 168(11):5659-5666.
37. Johnston CM, Wood AL, Bolland DJ et al. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. J Immunol 2006; 176(7):4221-4234.
38. Yancopoulos GD, Desiderio SV, Paskind M et al. Preferential utilization of the most JH-proximal VH gene segments in pre-B-cell lines. Nature 1984; 311(5988):727-733.
39. Jeong HD, Teale JM. VH gene family repertoire of resting B-cells. Preferential use of D-proximal families early in development may be due to distinct B-cell subsets. J Immunol 1989; 143(8):2752-2760.
40. Malynn BA, Yancopoulos GD, Barth JE et al. Biased expression of JH-proximal VH genes occurs in the newly generated repertoire of neonatal and adult mice. J Exp Med 1990; 171(3):843-859.
41. Ten Boekel E, Melchers F, Rolink AG. Changes in the V(H) gene repertoire of developing precursor B-lymphocytes in mouse bone marrow mediated by the pre-B-cell receptor. Immunity 1997; 7(3):357-368.
42. Corcoran AE, Riddell A, Krooshoop D et al. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. Nature 1998; 391(6670):904-907.
43. Bertolino E, Reddy K, Medina KL et al. Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. Nat Immunol 2005; 6(8):836-843.
44. Johnson K, Angelin-Duclos C, Park S et al. Changes in histone acetylation are associated with differences in accessibility of V(H) gene segments to V-DJ recombination during B-cell ontogeny and development. Mol Cell Biol 2003; 23(7):2438-2450.
45. Nutt SL, Heavey B, Rolink AG et al. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature 1999; 401(6753):556-562.
46. Hesslein DG, Pflugh DL, Chowdhury D et al. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. Genes Dev 2003; 17(1):37-42.
47. Johnson K, Pflugh DL, Yu D et al. B-cell-specific loss of histone 3 lysine 9 methylation in the V(H) locus depends on Pax5. Nat Immunol 2004; 5(8):853-861.
48. Hsu LY, Liang HE, Johnson K et al. Pax5 activates immunoglobulin heavy chain V to DJ rearrangement in transgenic thymocytes. J Exp Med 2004; 199(6):825-830.
49. Liu H, Schmidt-Supprian M, Shi Y et al. Yin Yang 1 is a critical regulator of B-cell development. Genes Dev 2007; 21(10):1179-1189.
50. Curry JD, Geier JK, Schlissel MS. Single-strand recombination signal sequence nicks in vivo: evidence for a capture model of synapsis. Nat Immunol 2005; 6(12):1272-1279.
51. Su IH, Basavaraj A, Krutchinsky AN et al. Ezh2 controls B-cell development through histone H3 methylation and Igh rearrangement. Nat Immunol 2003; 4(2):124-131.
52. Osipovich O, Milley R, Meade A et al. Targeted inhibition of V(D)J recombination by a histone methyltransferase. Nat Immunol 2004; 5(3):309-316.
53. Bolland DJ, Wood AL, Johnston CM et al. Antisense intergenic transcription in V(D)J recombination. Nat Immunol 2004; 5(6):630-637.
54. Jhunjhunwala S, van Zelm MC, Peak MM et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell 2008; 133(2):265-279.
55. Rolink AG, Winkler T, Melchers F et al. Precursor B-cell receptor-dependent B-cell proliferation and differentiation does not require the bone marrow or fetal liver environment. J Exp Med 2000; 191(1):23-32.
56. Tokoyoda K, Egawa T, Sugiyama T et al. Cellular niches controlling B-lymphocyte behavior within bone marrow during development. Immunity 2004; 20(6):707-718.
57. McGuire KL, Vitetta ES. kappa/lambda Shifts do not occur during maturation of murine B-cells. J Immunol 1981; 127(4):1670-1673.

58. Johnson K, Hashimshony T, Sawai CM et al. Regulation of Immunoglobulin Light-Chain Recombination by the Transcription Factor IRF-4 and the Attenuation of Interleukin-7 Signaling. Immunity 2008; 28(3):335-345.
59. Lazorchak AS, Schlissel MS, Zhuang Y. E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B-cells. Mol Cell Biol 2006; 26(3):810-821.
60. Lu R, Medina KL, Lancki DW et al. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. Genes Dev 2003; 17(14):1703-1708.
61. Sato H, Saito-Ohara F, Inazawa J et al. Pax-5 is essential for kappa sterile transcription during Ig kappa chain gene rearrangement. J Immunol 2004; 172(8):4858-4865.
62. Shaffer AL, Peng A, Schlissel MS. In vivo occupancy of the kappa light chain enhancers in primary pro- and pre-B-cells: a model for kappa locus activation. Immunity 1997; 6(2):131-143.
63. Muljo SA, Schlissel MS. A small molecule Abl kinase inhibitor induces differentiation of Abelson virus-transformed pre-B-cell lines. Nat Immunol 2003; 4(1):31-37.
64. Thompson EC, Cobb BS, Sabbattini P et al. Ikaros DNA-binding proteins as integral components of B-cell developmental-stage-specific regulatory circuits. Immunity 2007; 26(3):335-344.
65. Spanopoulou E, Roman CA, Corcoran LM et al. Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. Genes Dev 1994; 8(9):1030-1042.
66. Young F, Ardman B, Shinkai Y et al. Influence of immunoglobulin heavy- and light-chain expression on B-cell differentiation. Genes Dev 1994; 8(9):1043-1057.
67. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85(6):887-897.
68. Shaw AC, Swat W, Davidson L et al. Induction of Ig light chain gene rearrangement in heavy chain-deficient B-cells by activated Ras. Proc Natl Acad Sci USA 1999; 96(5):2239-2243.
69. Flemming A, Brummer T, Reth M et al. The adaptor protein SLP-65 acts as a tumor suppressor that limits pre-B-cell expansion. Nat Immunol 2003; 4(1):38-43.
70. Xu S, Lee KG, Huo J et al. Combined deficiencies in Bruton tyrosine kinase and phospholipase Cgamma2 arrest B-cell development at a pre-BCR+ stage. Blood 2007; 109(8):3377-3384.
71. Ma S, Turetsky A, Trinh L et al. IFN regulatory factor 4 and 8 promote Ig light chain kappa locus activation in pre-B-cell development. J Immunol 2006; 177(11):7898-7904.
72. Grawunder U, Haasner D, Melchers F et al. Rearrangement and expression of kappa light chain genes can occur without mu heavy chain expression during differentiation of pre-B-cells. Int Immunol 1993; 5(12):1609-1618.
73. Milne CD, Fleming HE, Paige CJ. IL-7 does not prevent pro-B/pre-B-cell maturation to the immature/sIgM(+) stage. Eur J Immunol 2004; 34(10):2647-2655.
74. Coleclough C, Perry RP, Karjalainen K et al. Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. Nature 1981; 290(5805):372-378.
75. Schlissel MS. Regulation of activation and recombination of the murine Igkappa locus. Immunol Rev 2004; 200:215-223.
76. Liang HE, Hsu LY, Cado D et al. Variegated transcriptional activation of the immunoglobulin kappa locus in pre-B-cells contributes to the allelic exclusion of light-chain expression. Cell 2004; 118(1):19-29.
77. Mostoslavsky R, Singh N, Tenzen T et al. Asynchronous replication and allelic exclusion in the immune system. Nature 2001; 414(6860):221-225.
78. Mostoslavsky R, Kirillov A, Ji YH et al. Demethylation and the establishment of kappa allelic exclusion. Cold Spring Harb Symp Quant Biol 1999; 64:197-206.
79. Goldmit M, Schlissel M, Cedar H et al. Differential accessibility at the kappa chain locus plays a role in allelic exclusion. EMBO J 2002; 21(19):5255-5261.
80. Goldmit M, Ji Y, Skok J et al. Epigenetic ontogeny of the Igk locus during B-cell development. Nat Immunol 2005; 6(2):198-203.
81. Liu Z, Widlak P, Zou Y et al. A recombination silencer that specifies heterochromatin positioning and ikaros association in the immunoglobulin kappa locus. Immunity 2006; 24(4):405-415.
82. Chowdhury D, Sen R. Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements. Immunity 2003; 18(2):229-241.
83. Roldan E, Fuxa M, Chong W et al. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol 2005; 6(1):31-41.
84. Hewitt SL, Farmer D, Marszalek K et al. Association between the Igk and Igh immunoglobulin loci mediated by the 3' Igk enhancer induces 'decontraction' of the Igh locus in pre-B-cells. Nat Immunol 2008; 9(4):396-404.

Chapter 12

Regulation of V(D)J Recombination by E-Protein Transcription Factors

Mary Elizabeth Jones and Yuan Zhuang*

Abstract

Extensive study of the E-proteins E2A and HEB during lymphocyte development has revealed various functions for these bHLH transcription factors in regulating V(D)J recombination in both B- and T-cells. The study of E-proteins in mammals began with the identification of E2A by its ability to bind immunoglobulin heavy and light chain enhancers. Subsequent analysis has identified numerous roles for E2A and HEB at the immunoglobulin and T-cell receptor loci. E-protein targets also include the *rag* genes and other factors critical for recombination and for regulation of the developmental windows when cells undergo recombination. E-proteins appear to be master regulators that coordinate antigen receptor gene rearrangement and expression. This chapter focuses on how E-proteins regulate V(D)J recombination by activating transcription, initiating rearrangement and driving differentiation during B- and T-cell development.

Introduction

E2A, the founding member of the E-protein family of transcription factors in mammals, was originally identified by its ability to bind enhancer regions of the immunoglobulin heavy chain (IgH) and light chain (IgL) genes. Early analysis of the IgH and IgL enhancers identified a conserved sequence that serves as a tissue-specific protein binding site in B-cells.[1,2] The two alternatively spliced products of the *e2a* gene, E47 and E12, were later isolated as the proteins binding to this conserved sequence, which is defined as an E-box site.[3,4] A much broader role for *E2A* in development was immediately predicted due to its structural homology to the Drosophila gene *daughterless (da)*, involved in cell determination and differentiation.[3,5] Following their identification, E2A and the additional members of the mammalian E-protein family, HEB and E2-2, have been extensively studied for their critical roles during lymphocyte development.[6]

E-proteins are basic helix-loop-helix (bHLH) transcription factors that function as dimers to bind DNA and regulate gene expression. The HLH region mediates protein dimerization and the basic region mediates DNA binding. E-protein dimers bind to E-box sites, defined by the consensus sequence CANNTG. E2A homodimers and E2A/HEB heterodimers are the primary E-protein dimers functioning in B- and T-cells, respectively. The DNA binding activity of E-protein dimers is negatively regulated by the four members of the Id (inhibitor of differentiation) protein family, Id1-Id4. Id proteins contain an HLH motif for dimerization but lack a DNA binding basic region, thus allowing competitive dimerization to inhibit E-protein activity. The balance of E-protein and Id expression is tightly regulated throughout B- and T-cell development.

Association of E2A with the Ig enhancers strongly suggests a role for E-proteins in regulating V(D)J recombination. E2A binds directly to E-box sites within the IgH Eμ enhancer and IgL kappa

*Corresponding Author: Yuan Zhuang—Department of Immunology, Duke University Medical Center, Box 3010, Durham, NC 27710, U.S.A. Email: yzhuang@duke.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

(Igk) intronic and 3' enhancers.[7-10] Additional regions within the Ig and T-cell receptor (TCR) loci also contain E-box sites. For example, putative E-box sites have been identified downstream of the recombination signal sequence (RSS) within most Igk V gene families[11] and the TCRα enhancer contains an E-box site with sequence similarity to the Ig enhancer site.[12] E-box sites are also located within the TCRβ enhancer.[13,14] Two E-box motifs are located in the core β enhancer region responsible for enhancer-dependent recombination activity, and nuclear factor binding has been suggested at one of these sites by DNA footprinting analysis.[15] In addition to sequence analysis of Ig and TCR regulatory regions, gene knockout and over-expression models have further suggested roles for E-proteins in V(D)J recombination during lymphocyte development. Accumulating evidence indicates multiple ways through which E-proteins directly or indirectly impact V(D)J recombination in both B- and T-cells. E-proteins can regulate V(D)J recombination at various levels, including the transcriptional control of Ig and TCR associated genes, initiation of gene rearrangement and regulation of differentiation through the developmental stages when Ig and TCR loci recombine.

Transcriptional Control of Ig and TCR Antigen Receptor and Their Associated Genes

E-proteins activate transcription of multiple factors essential for V(D)J recombination, including the Ig and TCR genes themselves. Sterile germline transcripts through Ig and TCR loci have been hypothesized to play a role in increasing chromatin accessibility prior to recombination.[16] An example of this role for transcription has recently been shown at the TCRα locus.[17] When transcription is blocked within the Jα locus, both rearrangement and chromatin remodeling are suppressed. It is therefore possible that E-proteins may be impacting chromatin accessibility and recombination through activation of germline transcription. There have been various examples of E2A inducing transcription within the Ig and TCR loci, mostly through in vitro studies in cell lines. Over-expression of E2A in nonB-cell lines is sufficient to induce ectopic expression of germline transcripts from the IgH and Igk loci. Forced expression of E47 has been shown to induce IgH transcription in pre-T and fibroblast cell lines.[18,19] E12 has been shown to induce Igk transcription in a mitogen stimulated macrophage cell line[20] and E12 or E47 can also activate Igk transcription in a kidney cell line.[21] Consistent with these results, Id over-expression in B-cell lines inhibits the activity of both IgH and Igk enhancers to induce transcription, indicating the role for E2A in activating enhancer-dependent transcription at these loci.[22] In addition, loss of E2A in pre-B-cell lines results in a loss of Igk transcription.[23] E2A may not only regulate Igk transcription through interaction with the intronic and 3′ enhancers, but may also function at the Igk promoters, where conserved E-box sites can also be found.[24]

A similar role for E-protein mediated transcriptional activation has been suggested for the TCR loci as well. Over-expression of E2A and/or HEB in a kidney cell line activates Vγ and Vδ germline transcription.[25] In this study, E2A and HEB activate only a specific subset of Vγ and Vδ genes and upon cotransfection with Ragl and Rag2, rearrangements utilizing these specific V segments are induced. This correlation suggests E-protein activation of transcription is linked to recombination at these loci. Putative E-box sites have been described within the Vβ promoter regions,[26] but whether or not E-proteins play a similar role in activating germline transcription at the TCRβ loci is still under investigation.

E-protein downstream targets relative to V(D)J recombination also include genes encoding the recombinase machinery and several receptor components that pair with the functionally rearranged Ig and TCR chains. Two of these targets most essential to V(D)J recombination are the recombination activating genes, *ragl* and *rag2*. *Ragl* expression is induced upon over-expression of E12 in a macrophage cell line and *Ragl* and *Rag2* expression levels increase upon over-expression of E47 in a pre-T-cell line.[19,20] E2A has also been implicated in regulating *Rag* expression by interacting with the Erag enhancer, critical for *Rag* expression in B-cells.[27] Forced expression of Id3 in T-cell progenitors inhibits the up-regulation of *Ragl* and *Rag2*, further demonstrating a role for E-proteins in initiation of *rag* gene expression.[28] Another E2A target

critical during V(D)J recombination is the gene encoding terminal deoxynucleotide transferase (TdT). E2A binding has been observed at the 5′ region of the *tdt* locus and E47 can activate *TdT* expression in a nonlymphoid cell line.[7,18]

Finally, E-proteins regulate components of both the pre-B-cell receptor (pre-BCR) and pre-TCR. E-proteins activate transcription of surrogate light chain genes (*λ5* and *Vpre B*) and *pre-Tα*, which are required to pair with IgH and TCRβ, respectively.[7,10,20,29-33] This pairing allows developing B-cells to express a pre-BCR and developing αβ T-cells to express a pre-TCR. E2A also regulates expression of mb-1 and possibly *B29*, additional components of the pre-BCR.[7,34,35] Surface expression of a pre-BCR or pre-TCR triggers entry to the next stage of development where the cells will then undergo rearrangement of IgL and TCRα genes, respectively. This role for E-proteins in regulating differentiation through the stages when recombination occurs will be further discussed in a later section of this chapter.

Induction of Ig and TCR Gene Rearrangement

Ectopic expression of E-proteins in nonlymphoid cells not only activates transcription, but also induces rearrangement events in the Ig and TCR loci upon co-expression with Rag1 and Rag2. As mentioned above, introduction of E2A and/or HEB with the Rag proteins in a kidney cell line induces rearrangements within the TCRγ and TCRδ loci.[25,36] In separate studies, transfection of E2A and Rag was shown to induce IgH D-J rearrangement in a pre-T-cell line and IgH D-J and Igk Vk1-J rearrangements in a kidney cell line.[19,21,37] In each of these cases, E2A generates a diverse repertoire, yet only certain subsets of gene segments are targeted for recombination. The mechanism by which E-proteins mediate recombination is not entirely understood. One possibility is that E-proteins create localized accessibility for recombination and therefore may influence the relative rearrangement efficiency of specific gene subsets.[37]

The physiological role of E2A in V(D)J recombination has been further defined by in vivo and in vitro studies of Igk rearrangement in B-cells. Targeted mutation of the two functional E-box sites within the Igk intronic enhancer results in a severe reduction in Igk rearrangement in developing B-cells[38] and deletion of E2A in pre-B-cell lines blocks Igk rearrangement.[23] In addition, re-introduction of E47 to these E2A deficient pre-B-cell lines rescues Igk recombination.[23] These studies suggest that E-proteins regulate initiation of V(D)J recombination at least in part by directly binding to cis-regulatory elements within the recombining loci.

E-proteins have also been proposed to regulate secondary IgL rearrangement in immature B-cells.[39] E2A wild-type mice expressing an auto-reactive BCR transgene display a significant population of peripheral B-cells that have undergone a secondary rearrangement of the endogenous IgL to replace the auto-reactive BCR. However, E2A heterozygous mice expressing the auto-reactive BCR transgene contain very few mature B-cells. This suggests that E2A dosage is critical for B-cells to undergo receptor editing, allowing replacement of an auto-reactive receptor.

Regulation of the Developmental Window for V(D)J Recombination

In addition to directly activating transcription and initiating rearrangement as described above, E-proteins also indirectly regulate V(D)J recombination by controlling differentiation during B- and T-cell development. Since E-proteins are expressed in both B- and T-cells, there are obviously additional factors determining the lineage and stage specific recombination events at the Ig and TCR loci. Failure of cells to enter the stage when these factors are functioning would prevent initiation of rearrangement events. Defects in Ig or TCR recombination in E-protein deficient models may often result from a block in development prior to the stage when cells would undergo rearrangement. For example, E2A deficient mice exhibit a block in B-cell development prior to IgH rearrangement.[40,41] E2A deficient B-cells are blocked at the prepro-B-cell stage, a stage prior to the pro-B-cell stage where IgH intronic enhancer deficient mice demonstrate a block.[42] This suggests that even though E2A has been shown to play a role in activating the IgH enhancer, E2A has additional roles prior to this role that contribute to the block in IgH recombination in E2A deficient mice. Even though many of the E2A targets at this early stage of B-cell development

remain unidentified, potential targets have been revealed through microarray analysis of E2A deficient cells and upon over expression of E2A in these cells.[10,43,44] The remaining challenge is to identify which of these targets are critical for E2A mediated development to the pro-B-cell stage for subsequent rearrangement of IgH genes.

Once developing B-cells have undergone IgH rearrangement, E2A remains critical for the expression of the surrogate light chain components.[7,10,20,30-32] Vpre-B- and λ5 are required for surface pre-BCR expression and proper differentiation to the pre-B-cell stage where the cells will undergo IgL recombination.[45] Although E2A is also critical during IgL rearrangement, E2A first regulates differentiation to the pre-B-cell stage. If E2A is required throughout development of pro and pre-B-cells, how does it regulate IgH and IgL in stage specific manners? Tissue and stage specific expression of factors that cooperate with E2A can result in activation of different sets of genes. For example, E2A cooperates with the B-cell specific transcription factors early B-cell factor (EBF) and Pax5 to regulate expression of *mb-1* in pre-B-cells.[35] The differential transcriptional networks established by E2A at the pro-B vs pre-B-cell stages could contribute to the stage specific effects of E2A at the IgH and IgL loci. Other potential mechanisms responsible for E-protein stage and lineage specific regulation of receptor gene loci will be discussed further in the final section of this chapter.

A similar role also exists for E2A and HEB during differentiation of developing T-cells. Since T- cell development is regulated by the combined dosage of E2A and HEB, single knockouts exhibit only partial blocks in T-cell development.[46,47] To inhibit total E-protein activity, mice expressing a dominant negative form of HEB were generated.[48] The dominant negative HEB protein is able to form nonfunctional heterodimers with E2A to inhibit both E2A and HEB activity, therefore resulting in a more severe phenotype than that seen in the single knockout mice. Dominant negative HEB mice exhibit a block in T-cell development at the $CD4^-CD8^-$ double negative (DN) stage and a defect in TCRβ V(D)J recombination. Introduction of a functional αβ TCR transgene is unable to rescue this developmental block, indicating that the rearrangement defect is not the only cause for the block at DN stage. These results demonstrate that E-proteins have multiple roles during this window of development. Since these roles include regulation of differentiation, *Rag* expression and perhaps TCRβ expression and rearrangement, it is likely that multiple E-protein targets are responsible for coordinating V(D)J recombination at this stage. E-proteins are then also required for progression from DN to DP, partly through the induction of *pre-Tα* expression.[49] E-proteins therefore regulate the entry and progression through stages critical for both TCRβ and TCRα recombination.

Proper regulation of gene segment usage during V(D)J recombination within the TCRγ and δ loci is also dependent on E-proteins. There is a differential usage of Vγ and Vδ genes during rearrangement in fetal vs. adult thymocyte development.[50] Adult E2A deficient mice display a defect in usage of adult predominant Vγ2 and Vδ5 genes whereas rearrangements utilizing the fetal specific Vγ3 and Vδ1 gene segments persist.[51] These results indicate that E2A positively and negatively regulates specific V genes during the window of adult T-cell development. This study also demonstrates a requirement for E2A during fetal thymocyte development for usage of a few V gene segments, but Vγ3 and Vδ1 fetal usage appears comparable to wild-type. The mechanism by which E2A activity results in the increased usage of some V genes and repression of others is not well understood. The mechanism by which E2A promotes usage of gene segments in adult but not fetal development is suspected to result from different dosages of E2A activity.[51,52] Even though *e2A* is expressed at comparable levels in both adult and fetal thymus, *Id2* expression is higher in fetal thymus, which would be expected to result in reduced E2A activity in fetal compared to adult thymus.[52]

Finally, accumulating data indicates that E-proteins can also influence the duration of the recombining window of development. An example of this role is seen at the TCRα locus during the $CD4^+CD8^+$ double positive (DP) stage. The transcription factor RORγt, an isoform of the orphan nuclear receptor RORγ, is required in DP thymocytes to regulate the survival window at this stage by inducing $Bcl\text{-}X_L$ expression.[53-55] This DP survival window is critical for establishing

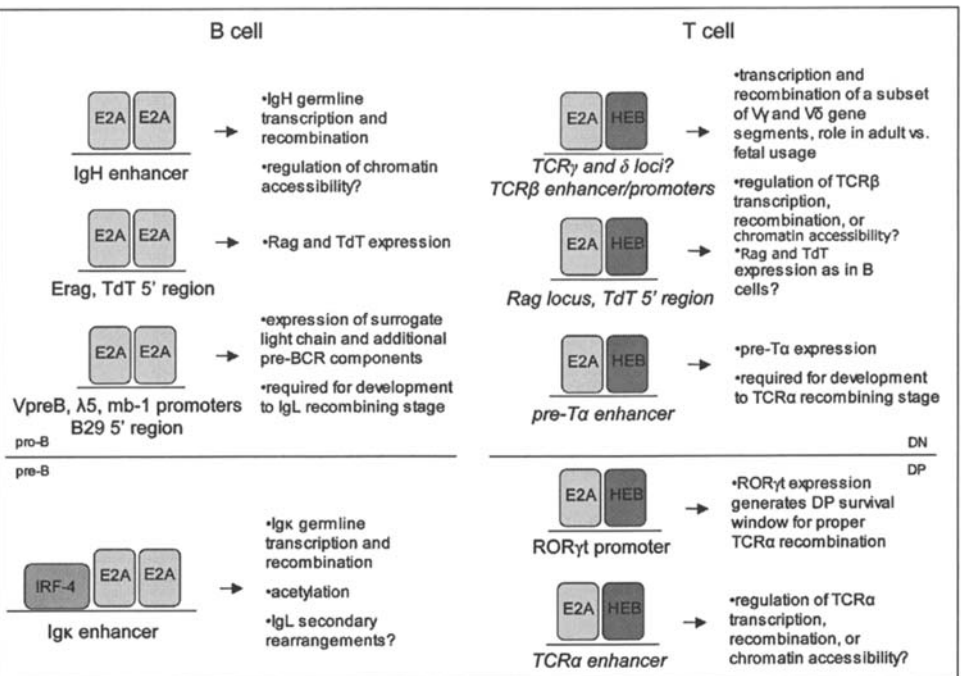


Figure 1. E-protein mediated regulation of V(D)J recombination during B- and T-cell development. E2A homodimers and E2A/HEB heterodimers are displayed for their roles relative to V(D)J recombination in B- vs T-cells, respectively. E-proteins regulate Ig and TCR loci, genes required for rearrangement and components of the pre-BCR and pre-TCR during proB and DN stages of B- and T-cell development, respectively. For B-cell and αβT-cell development, each of these events is critical for cells to differentiate to pre-B- and DP stages to then undergo IgL and TCRα recombination. E-proteins are not only required for differentiation to these stages, but are then further required for proper IgL and TCRα recombination. E-proteins are also critical for proper V gene usage during fetal and adult γδT-cell development. Depicted examples here are limited to known or predicted sites of E-protein DNA-binding within target genes. Potential E-protein binding sites which have not been demonstrated as direct binding sites in the specified cell type by ChIP analysis are in italics. E-proteins may also serve functions independent of DNA-binding, which are not included in this figure.

a diverse repertoire of TCRα rearrangements. Since rearrangements through the Jα locus during TCRα recombination occur in a proximal to distal manner, 5′ to 3′, the lifespan of DP cells can influence the repertoire.[56-59] *RORγt* deficient mice, exhibiting a shorter DP lifespan, also exhibit a defect in usage of 3′ Jα gene segments.[59] E2A has been shown to activate expression of *RORγt* in thymocytes by binding to critical E-box sites within the promoter region.[53] In agreement with these findings, a recent study demonstrates a similar 5′ skewing of Jα usage when both E2A and HEB are deleted at the DP stage (ME Jones and Y Zhuang data to be published). Likely through regulation of RORγt expression, E2A and HEB indirectly influence the TCRα repertoire by ensuring a sufficient window for rearrangement.

Conclusion

E-proteins demonstrate considerable involvement in various aspects of V(D)J recombination, a few of which are depicted in Figure 1. Two main questions remain. First, what are the underlying mechanisms guiding E-protein mediated transcriptional regulation in a lineage and stage specific fashion? Second, what roles are E-proteins playing in addition to acting as transcriptional regulators? As mentioned earlier, E-proteins are suspected to generate localized accessibility around specific gene segments within various receptor loci.[37] E-proteins have also been suggested by additional studies to play a role in chromatin modification. Cooperative efforts of E2A, EBF and Pax5 have been shown to regulate CpG demethylation and nucleosome remodeling at the *mb-1* promoter.[34] More relative to V(D)J recombination, E2A has also been shown to play a role in Igk enhancer acetylation.[23] If E-proteins can induce chromatin accessibility for recombination in certain localized regions within both the Ig and TCR loci, how do E-proteins regulate Ig and TCR receptors specifically in B- and T-cells, respectively? Even though expression of E2A in nonlymphoid cells can induce rearrangements, it is important to remember that these are mostly over-expression studies and E2A may be inducing expression of additional factors that are repressed in B- or T-cells. These results suggest that overall E-protein dosage may play a role in differential gene activation. For example, some targets may require a certain threshold of E-protein activity to be activated. This threshold would be expected to be exceeded in over-expression studies, but may be differentially regulated in B- and T-cells. Also, limited access to E-box sites in B- vs T-cells could potentially contribute to E-proteins' B- vs T-cell specific effects.

Another way E-proteins could be exhibiting lineage and stage specific affects is through regulated interactions with different binding factors. So far only a few co-activators interacting with E-proteins in lymphocytes have been identified. One group of factors that have been shown to associate with E-proteins are the histone acetyltransferases (HATs) p300, CBP and PCAF.[60-63] One study shows these interactions existing in B-cells and demonstrates that HATs can enhance E2A transcriptional activity.[62] However, which E2A target genes are dependent on E2A-HAT interactions have yet to be determined. The corepressor ETO has also been shown to interact with E-proteins and in doing so, blocks the recruitment of HATs.[64] ETO is also able to bind to histone deacetylases (HDACs).[64,65] The ability of E-proteins to recruit either HATs or HDACs could potentially contribute to the lineage and stage specific effects of E-proteins at the Ig and TCR loci. Another way E-proteins could have lineage and stage specific functions is through recruitment of E-proteins by factors with more restricted expression patterns. For example, IRF-4 has been shown to promote E2A recruitment at the Igk 3′ enhancer in pre-B-cells.[23] Future studies will likely shed more light on how the somewhat ubiquitous, yet tightly regulated, expression of E-proteins can result in lineage and stage specific regulation of the Ig and TCR genes.

References

1. Ephrussi A, Church GM, Tonegawa S et al. B lineage—specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science 1985; 227(4683):134-140.
2. Church GM, Ephrussi A, Gilbert W et al. Cell-type-specific contacts to immunoglobulin enhancers in nuclei. Nature 1985; 313(6005):798-801.
3. Murre C, McCaw PS, Vaessin H et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 1989; 58(3):537-544.

4. Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD and myc proteins. Cell 1989; 56(5):777-783.
5. Jan YN, Jan LY. HLH proteins, fly neurogenesis and vertebrate myogenesis. Cell 1993; 75(5):827-830.
6. Murre C. Helix-loop-helix proteins and lymphocyte development. Nat Immunol 2005; 6(11):1079-1086.
7. Greenbaum S, Zhuang Y. Identification of E2A target genes in B-lymphocyte development by using a gene tagging-based chromatin immunoprecipitation system. Proc Natl Acad Sci USA 2002; 99(23):15030-15035.
8. Murre C, Voronova A, Baltimore D. B-cell- and myocyte-specific E2-box-binding factors contain E12/E47-like subunits. Mol Cell Biol 1991; 11(2):1156-1160.
9. Henthorn P, Kiledjian M, Kadesch T. Two distinct transcription factors that bind the immunoglobulin enhancer microE5/kappa 2 motif. Science 1990; 247(4941):467-470.
10. Greenbaum S, Lazorchak AS, Zhuang Y. Differential functions for the transcription factor E2A in positive and negative gene regulation in pre-B-lymphocytes. J Biol Chem 2004; 279(43):45028-45035.
11. Brekke KM, Garrard WT. Assembly and analysis of the mouse immunoglobulin kappa gene sequence. Immunogenetics 2004; 56(7):490-505.
12. Ho IC, Yang LH, Morle G et al. A T-cell-specific transcriptional enhancer element 3' of C alpha in the human T-cell receptor alpha locus. Proc Natl Acad Sci USA 1989; 86(17):6714-6718.
13. Takeda J, Cheng A, Mauxion F et al. Functional analysis of the murine T-cell receptor beta enhancer and characteristics of its DNA-binding proteins. Mol Cell Biol 1990; 10(10):5027-5035.
14. Gottschalk LR, Leiden JM. Identification and functional characterization of the human T-cell receptor beta gene transcriptional enhancer: common nuclear proteins interact with the transcriptional regulatory elements of the T-cell receptor alpha and beta genes. Mol Cell Biol 1990; 10(10):5486-5495.
15. Tripathi RK, Mathieu N, Spicuglia S et al. Definition of a T-cell receptor beta gene core enhancer of V(D)J recombination by transgenic mapping. Mol Cell Biol 2000; 20(1):42-53.
16. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4(7):624-630.
17. Abarrategui I, Krangel MS. Regulation of T-cell receptor-alpha gene recombination by transcription. Nat Immunol 2006; 7(10):1109-1115.
18. Choi JK, Shen CP, Radomska HS et al. E47 activates the Ig-heavy chain and TdT loci in non-B-cells. EMBO J 1996; 15(18):5014-5021.
19. Schlissel M, Voronova A, Baltimore D. Helix-loop-helix transcription factor E47 activates germ-line immunoglobulin heavy-chain gene transcription and rearrangement in a pre-T-cell line. Genes Dev 1991; 5(8):1367-1376.
20. Kee BL, Murre C. Induction of early B-cell factor (EBF) and multiple B lineage genes by the basic helix-loop-helix transcription factor E12. J Exp Med 1998; 188(4):699-713.
21. Romanow WJ, Langerak AW, Goebel P et al. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Mol Cell 2000; 5(2):343-353.
22. Wilson RB, Kiledjian M, Shen CP et al. Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: implications for B-lymphoid-cell development. Mol Cell Biol 1991; 11(12):6185-6191.
23. Lazorchak AS, Schlissel MS, Zhuang Y. E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B-cells. Mol Cell Biol 2006; 26(3):810-821.
24. Bemark M, Liberg D, Leanderson T. Conserved sequence elements in K promoters from mice and humans: implications for transcriptional regulation and repertoire expression. Immunogenetics 1998; 47(3):183-195.
25. Ghosh JK, Romanow WJ, Murre C. Induction of a diverse T-cell receptor gamma/delta repertoire by the helix-loop-helix proteins E2A and HEB in nonlymphoid cells J Exp Med 2001; 193(6):769-776.
26. Chen F, Rowen L, Hood L et al. Differential transcriptional regulation of individual TCR V beta segments before gene rearrangement. J Immunol 2001; 166(3):1771-1780.
27. Hsu LY, Lauring J, Liang HE et al. A conserved transcriptional enhancer regulates RAG gene expression in developing B-cells. Immunity 2003; 19(1):105-117.
28. Blom B, Heemskerk MH, Verschuren MC et al. Disruption of alpha beta but not of gamma delta T-cell development by overexpression of the helix-loop-helix protein Id3 in committed T-cell progenitors. EMBO J 1999; 18(10):2793-2802.
29. Herblot S, Steff AM, Hugo P et al. SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression. Nat Immunol 2000; 1(2):138-144.
30. Ikawa T, Kawamoto H, Wright LY et al. Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. Immunity 2004; 20(3):349-360.
31. Sigvardsson M. Overlapping expression of early B-cell factor and basic helix-loop-helix proteins as a mechanism to dictate B-lineage-specific activity of the lambda5 promoter. Mol Cell Biol 2000; 20(10):3640-3654.

32. Sigvardsson M, O'Riordan M, Grosschedl R. EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. Immunity 1997; 7(1):25-36.
33. Reizis B, Leder P. Expression of the mouse pre-T-cell receptor alpha gene is controlled by an upstream region containing a transcriptional enhancer. J Exp Med 1999; 189(10):1669-1678.
34. Maier H, Ostraat R, Gao H et al. Early B-cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. Nat Immunol 2004; 5(10):1069-1077.
35. Sigvardsson M, Clark DR, Fitzsimmons D et al. Early B-cell factor, E2A and Pax-5 cooperate to activate the early B-cell-specific mb-1 promoter. Mol Cell Biol 2002; 22(24):8539-8551.
36. Langerak AW, Wolvers-Tettero IL, van Gastel-Mol EJ et al. Basic helix-loop-helix proteins E2A and HEB induce immature T-cell receptor rearrangements in nonlymphoid cells. Blood 2001; 98(8):2456-2465.
37. Goebel P, Janney N, Valenzuela JR et al. Localized gene-specific induction of accessibility to V(D)J recombination induced by E2A and early B-cell factor in nonlymphoid cells. J Exp Med 2001; 194(5):645-656.
38. Inlay MA, Tian H, Lin T et al. Important roles for E protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating Vkappa Jkappa rearrangement. J Exp Med 2004; 200(9):1205-1211.
39. Quong MW, Martensson A, Langerak AW et al. Receptor editing and marginal zone B-cell development are regulated by the helix-loop-helix protein, E2A. J Exp Med 2004; 199(8):1101-1112.
40. Bain G, Maandag EC, Izon DJ et al. E2A proteins are required for proper B-cell development and initiation of immunoglobulin gene rearrangements. Cell 1994; 79(5):885-892.
41. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B-cell formation. Cell 1994; 79(5):875-884.
42. Perlot T, Alt FW, Bassing CH et al. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci USA 2005; 102(40):14362-14367.
43. Schwartz R, Engel I, Fallahi-Sichani M et al. Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling and T lineage development. Proc Natl Acad Sci USA 2006; 103(26):9976-9981.
44. Ikawa T, Kawamoto H, Goldrath AW et al. E proteins and Notch signaling cooperate to promote T-cell lineage specification and commitment. J Exp Med 2006; 203(5):1329-1342.
45. Shimizu T, Mundt C, Licence S et al. Vpre-B1/Vpre-B2/lambda 5 triple-deficient mice show impaired B-cell development but functional allelic exclusion of the IgH locus. J Immunol 2002; 168(12):6286-6293.
46. Bain G, Engel I, Robanus Maandag EC et al. E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol 1997; 17(8):4782-4791.
47. Barndt R, Dai MF, Zhuang Y. A novel role for HEB downstream or parallel to the pre-TCR signaling pathway during alpha beta thymopoiesis. J Immunol 1999; 163(6):3331-3343.
48. Barndt RJ, Dai M, Zhuang Y. Functions of E2A-HEB heterodimers in T-cell development revealed by a dominant negative mutation of HEB. Mol Cell Biol 2000; 20(18):6677-6685.
49. Wojciechowski J, Lai A, Kondo M et al. E2A and HEB Are Required to Block Thymocyte Proliferation Prior to Pre-TCR Expression. J Immunol 2007; 178(9):5717-5726.
50. Raulet DH. The structure, function and molecular genetics of the gamma/delta T-cell receptor. Annu Rev Immunol 1989; 7:175-207.
51. Bain G, Romanow WJ, Albers K et al. Positive and negative regulation of V(D)J recombination by the E2A proteins. J Exp Med 1999; 189(2):289-300.
52. David-Fung ES, Yui MA, Morales M et al. Progression of regulatory gene expression states in fetal and adult pro-T-cell development. Immunol Rev 2006; 209:212-236.
53. Xi H, Schwartz R, Engel I et al. Interplay between RORgammat, Egr3 and E proteins controls proliferation in response to pre-TCR signals. Immunity 2006; 24(6):813-826.
54. Sun Z, Unutmaz D, Zou YR et al. Requirement for RORgamma in thymocyte survival and lymphoid organ development. Science 2000; 288(5475):2369-2373.
55. Kurebayashi S, Ueda E, Sakaue M et al. Retinoid-related orphan receptor gamma (RORgamma) is essential for lymphoid organogenesis and controls apoptosis during thymopoiesis. Proc Natl Acad Sci USA 2000; 97(18):10132-10137.
56. Wang F, Huang CY, Kanagawa O. Rapid deletion of rearranged T-cell antigen receptor (TCR) Valpha-Jalpha segment by secondary rearrangement in the thymus: role of continuous rearrangement of TCR alpha chain gene and positive selection in the T-cell repertoire formation. Proc Natl Acad Sci USA 1998; 95(20):11834-11839.
57. Thompson SD, Pelkonen J, Hurwitz JL. First T-cell receptor alpha gene rearrangements during T-cell ontogeny skew to the 5′ region of the J alpha locus. J Immunol 1990; 145(7):2347-2352.
58. Petrie HT, Livak F, Burtrum D et al. T-cell receptor gene recombination patterns and mechanisms: cell death, rescue and T-cell production. J Exp Med 1995; 182(1):121-127.

59. Guo J, Hawwari A, Li H et al. Regulation of the TCRalpha repertoire by the survival window of CD4(+) CD8(+) thymocytes. Nat Immunol 2002; 3(5):469-476.
60. Eckner R, Yao TP, Oldread E et al. Interaction and functional collaboration of p300/CBP and bHLH proteins in muscle and B-cell differentiation. Genes Dev 1996; 10(19):2478-2490.
61. Qiu Y, Sharma A, Stein R. p300 mediates transcriptional stimulation by the basic helix-loop-helix activators of the insulin gene. Mol Cell Biol 1998; 18(5):2957-2964.
62. Bradney C, Hjelmeland M, Komatsu Y et al. Regulation of E2A activities by histone acetyltransferases in B-lymphocyte development. J Biol Chem 2003; 278(4):2370-2376.
63. Massari ME, Grant PA, Pray-Grant MG et al. A conserved motif present in a class of helix-loop-helix proteins activates transcription by direct recruitment of the SAGA complex. Mol Cell 1999; 4(1):63-73.
64. Zhang J, Kalkum M, Yamamura S et al. E protein silencing by the leukemogenic AML1-ETO fusion protein. Science 2004; 305(5688):1286-1289.
65. Hug BA, Lazar MA. ETO interacting proteins. Oncogene 2004; 23(24):4270-4274.

Chapter 13

Temporal and Spatial Regulation of V(D)J Recombination: Interactions of Extrinsic Factors with the RAG Complex

Yun Liu, Li Zhang and Stephen Desiderio*

Abstract

In the course of lymphoid development, V(D)J recombination is subject to stringent locus-specific and temporal regulation. These constraints are ultimately responsible for several features peculiar to lymphoid development, including the lineage specificity of antigen receptor assembly, allelic exclusion and receptor editing. In addition, cell cycle phase-dependent regulation of V(D)J recombinase activity ensures that DNA rearrangement is completed by the appropriate mechanism of DNA repair. Regulation of V(D)J recombination involves interactions between the V(D)J recombinase—a heteromeric complex consisting of RAG-1 and RAG-2 subunits—and macromolecular assemblies extrinsic to the recombinase. This chapter will focus on those features of the recombinase itself—and in particular the RAG-2 subunit—that interact with extrinsic factors to establish patterns of temporal control and locus specificity in developing lymphocytes.

Functional Organization of RAG-1 and RAG-2

RAG-1 and RAG-2 are 1040 and 527 amino acid residues long, respectively. Residues 384 through 1008 of RAG-1 constitute the core fragment, which contains the catalytic site for DNA cleavage,[1-3] mediates binding to recombination signal sequences (RSSs)[4-6] and makes contacts with the coding flanks.[7,8] The core RAG-2 fragment (Fig. 1), consisting of residues 1 through 387, extends interactions of RAG-1 with the RSS and is essential for helical distortion near the scissile bond, a possible prerequisite for transesterification.[4-6,9] Accordingly, mutations that impair recombinase-mediated cleavage and joining have been identified in core RAG-2.[10]

Residues 387 through 527 of RAG-2 comprise the non-core region (Fig. 1) and are dispensable for DNA cleavage by the RAG proteins in vitro. Nonetheless, removal of this region reduces the efficiency of extrachromosomal recombination,[11-16] increases production of hybrid joints,[17] impedes endogenous V_H-to-DJ_H joining[12,18,19] and promotes aberrant recombination.[20] The mechanisms underlying these effects may be complex, as the non-core region includes multiple functional domains (Fig. 1B).

*Corresponding Author: Stephen Desiderio—Department of Molecular Biology and Genetics and Institute for Cell Engineering, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. Email: sdesider@jhmi.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

Temporal Regulation of V(D)J Recombination through Interactions with the RAG-2 Non-Core Region

The non-core region of RAG-2 supports the periodic destruction of RAG-2 protein. RAG-2 accumulates in quiescent cells and in dividing cells during the G1 phase; rapid degradation of RAG-2 begins at the G1-to-S transition and continues until the following entry into G1.[21-23] Consequently, the appearance of recombination signal end intermediates[24,25] and RAG-signal end complexes[26] is restricted to G0/G1. Destruction of RAG-2 is triggered by phosphorylation of threonine 490, which lies within a phylogenetically conserved cyclin-dependent kinase (Cdk) target site and is also dependent on a lysine-rich interval spanning amino acid residues 499-508.[21] Overlapping the RAG-2 degradation domain (Fig. 1B) is a noncanonical nuclear localization sequence that supports binding of importin 5 and nuclear import of RAG-2.[27] At the G1-to-S transition, phosphorylation of RAG-2 by cyclinA/Cdk2 permits association of RAG-2 with the Skp2-SCF ubiquitin ligase. This phosphorylation-dependent interaction is mediated by the F-box protein Skp-2 and its associated protein Cks1. Upon polyubiquitylation of RAG-2 by Skp2-SCF, RAG-2 is subjected to proteasomal degradation.[28]

The cell cycle dependence of V(D)J recombination may play a role in the coupling of DNA cleavage by the RAG complex to DNA repair. V(D)J recombination is normally completed by a form of DNA repair termed nonhomologous end joining (NHEJ). NHEJ is active throughout the cell cycle, but an alternative mechanism for double-strand DNA repair, homologous recombination (HR), is nearly inactive during G1.[29] In thymocytes of mice expressing RAG-2(T490A), aberrant recombinants resembling products of abortive homologous recombination are observed to accumulate.[21] These observations suggest that restriction of RAG-2 accumulation to the G0 and G1 cell cycle phases promotes the correct repair of V(D)J recombination intermediates by NHEJ, perhaps by temporal sequestration of RAG activity from HR.

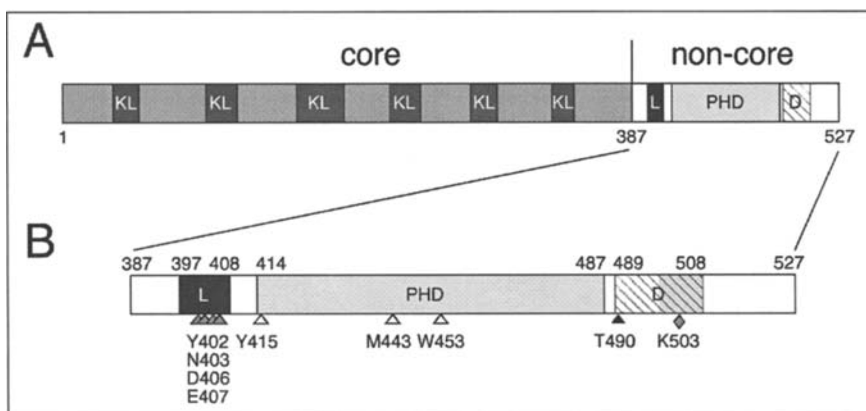


Figure 1. Regulatory domains of RAG-2. A) Schematic representation of mouse RAG-2. Core and non-core regions are designated; amino acid residues are numbered below. KL, Kelch-like propeller domains; L, linker domain; PHD, plant homeodomain finger; D, domain governing programmed degradation and nuclear import of RAG-2. B) Detailed representation of the non-core region. Amino acid residues at domain boundaries are numbered above. L (black rectangle), PHD (gray rectangle) and D (hatched rectangle) as defined in (A). The hatched interval denotes the extent of the domain governing cell cycle-dependent degradation of RAG-2; the shaded region within this interval marks the nuclear import signal that resides within the degradation domain. Shaded arrowheads, sites of mutations in the linker domain that impair V(D)J recombination. Open arrowheads, targets of mutations in the PHD domain that abolish H3K4me3 binding and impair V(D)J recombination. Black arrowhead, cyclinA/CDK2 phosphorylation site, essential for programmed degradation of RAG-2 at the G1-S transition. Shaded diamond, target of mutation that selectively impairs nuclear import of RAG-2.

Locus Specificity: General Remarks

The V(D)J recombinase is directed toward particular sets of gene segments, depending on lymphoid lineage and developmental stage. Recent work has begun to provide a framework for understanding how this targeting is achieved. At the level of unchromatinized DNA, the V(D)J recombinase is targeted to antigen receptor gene segments by means of specific interactions with flanking RSSs and this recognition does not require the non-core regions of RAG-1 or RAG-2. Not all RSSs support recombination with the same efficiency, because RSSs exhibit considerable sequence variation. Although sequence variation among RSSs can indeed affect gene segment usage,[30] these differences cannot account for the dynamic shifts in locus specificity that accompany commitment to distinct lymphoid lineages and developmental transitions within lineages. Rather, ordered rearrangement of antigen receptor gene segments is associated with the imposition or relief of epigenetic marks. Specific chromatin modifications in the vicinity of RSSs are strongly associated with the presence or absence of ongoing rearrangement. The propensity of a particular locus to undergo rearrangement has been thought to be determined by accessibility to the RAG complex, a view that ascribes a passive role to the recombinase. Recent findings, however, indicate that the recombinase—through direct binding to modified chromatin—is an active partner in the epigenetic regulation of rearrangement. We discuss below how epigenetic marks interact with the V(D)J recombinase to promote locus-specific rearrangement.

Epigenetic Modifications of Possible Relevance to V(D)J Recombination

An alteration in gene function is termed epigenetic if it is maintained through cell division and does not involve a change in the DNA sequence. One extensively studied epigenetic mark is DNA methylation on cytosine, which in mammals occurs at most CpG dinucleotides. A far more complex set of epigenetic marks are associated with the protein components of chromatin. The basic unit of eukaryotic chromatin is the nucleosome. This consists of a histone core—two molecules each of the histones H2A, H2B, H3 and H4—around which are wrapped about 146 base pairs of DNA. Histones are subject to a variety of posttranslational modifications including acetylation, methylation, phosphorylation, ubiquitylation and sumoylation. Differences in the degree and stereospecificity of modification contribute substantially to the complexity of these marks. Lysine, for example, can be mono-, di- or trimethylated, while arginine can be dimethylated symmetrically or asymmetrically. In addition to chemical modification, the register in which DNA is wrapped around the histone core—termed nucleosome phasing—can have profound effects on the accessibility of specific sequences to interacting factors. Observations relating these modes of epigenetic regulation to the activation or suppression of V(D)J recombination are summarized in turn below.

DNA Methylation

Methylation of CpG dinucleotides is normally associated with the suppression of transcription. Consistent with a general correlation of recombination with transcription, CpG methylation over antigen-receptor-gene segments is also associated with suppression of V(D)J recombination.[31] Deletion of PDβ1, a promoter located 5′ to the Dβ1 gene segment or Eβ an enhancer located 3′ to the TCRβ locus, is accompanied by increased CpG methylation in the Dβ1-Jβ1 region and defects in TCRβ1 rearrangement.[32-34] Conversely, demethylation of DNA has been associated with activation of rearrangement. In developing B-cells, for example, the Ig κ allele that is first activated for rearrangement is demethylated over the Jκ-Cκ region, while the opposite allele remains hypermethylated and is recruited to heterochromatin.[35,36]

Nucleosome Phasing

Together, the core RAG-1 and RAG-2 fragments catalyze RSS-specific nicking and transesterification of DNA substrates in vitro. Efficient cleavage is not observed, however, when chromatinized nuclear substrates are used.[37] RAG-mediated DNA cleavage in vitro is impeded when the target RSS is incorporated into a nucleosome;[38-40] the degree of inhibition has been variously proposed

to be dependent[38] or independent[39] of nucleosome phasing relative to the RSS. The resistance of mononucleosomal substrates to cleavage may result from inaccessibility of histone-associated DNA to the RAG complex as well as from helical distortion induced by wrapping of the DNA around the histone core.[40,41] The impediment to RAG-mediated DNA cleavage observed with mononucleosomal substrates in vitro can be relieved synergistically by histone acetylation and SWI/SNF-dependent remodeling, possibly as a result of alterations in chromatin structure that enhance accessibility of the RSS to the RAG complex.[40,41]

Histone Acetylation

Acetylation of histones H3 and H4 is associated with active chromatin. A positive correlation between histone acetylation and active antigen receptor gene rearrangement has been widely documented. Decreased acetylation of H3 and H4 is associated with diminished germline transcription at unrearranged antigen receptor loci and is important for allelic exclusion.[36,42-45] During B-cell development, diminished IL-7 signaling is associated with decreased histone acetylation and reduced accessibility to nucleases over distal V_H segments.[43] A similar relationship is observed over Vβ segments during the transition of intrathymic T-cell progenitors from the $CD4^-CD8^-$ to the $CD4^+CD8^+$ stage.[46] Thus, decreases in histone acetylation are associated with diminished rearrangement. Consistent with this relationship, Ig κ alleles at which recombination is active exhibit increased acetylation of histone H3.[36]

Histone H3 K9 Methylation

Dimethylation of histone H3 at lysine 9 (H3K9me2), which is associated with silent chromatin, is positively correlated with inhibition of V(D)J recombination.[47-49] Dimethyl marks at H3K9 are removed over V_H segments at the pro-B to pre-B-cell transition, at which stage V_H-to-DJ_H joining occurs; H3K9 demethylation is dependent on expression of the transcription factor Pax5 in pro-B-cells.[47] A role for H3K9me2 in the control of V(D)J recombination was suggested in an experiment that targeted G9a, a histone H3K9 methyltransferase, to a TCRβ minilocus. In this setting, directed H3K9 methylation was found to inhibit both germline transcription and V(D)J recombination, overriding the presence of cis-acting accessibility control elements.[49] An interpretation of these findings is complicated, because ablation of the G9a methyltransferase in mice had no significant effects on lymphoid development or stage specificity of V(D)J recombination, despite suppressive effects on λ light chain usage, B-cell proliferation and plasma cell differentiation.[50]

Histone H3 K4 Methylation

Methylation of histone H3 lysine 4 (H3K4) is a phylogenetically conserved modification that has been linked to transcriptional activation in yeast and metazoans.[51] The relationship between histone H3K4 methylation and V(D)J recombination has been the subject of much recent study.[36,48,52,53] Dimethylated histone H3K4 (H3K4me2)[48,53] and trimethylated H3K4 (H3K4me3)[54,55] exhibit distinct patterns of enhancement within the D-J_H cluster in pro-B-cells poised to undergo D-to-J_H rearrangement. Moreover, the recombinationally active Ig κ allele in pre-B-cells is marked by hypermethylation of H3K4.[36]

Monoubiquitylation of histone H2B at lysine 123 (ubH2B) promotes histone H3K4 methylation in yeast.[56-58] UbH2B is associated with transcriptionally active chromatin both in yeast[59-63] and in mammalian cells.[60,64] Patterns of ubH2B deposition have yet to be extensively mapped. As H2B ubiquitylation appears to be a prerequisite for H3K4 hypermethylation, it will be of interest to know whether the density of ubH2B is enhanced at sites of active V(D)J recombination, possibly extending the chain of causation one step upstream.

Direct Recognition of Modified Histone H3 by the V(D)J Recombinase

The observations outlined above, while essential to an understanding of epigenetic control, do not in themselves provide mechanistic insight into how histone modification is linked

mechanistically to V(D)J recombination. Building on recent progress in the understanding of how histone methylation patterns are read, several studies have combined biochemical, structural and genetic approaches to outline how one such linkage is established.

A variety of protein domains are capable of binding the N-terminal region of histone H3 when this is hypermethylated at lysine 4. These include the chromodomains of CHD1,[65,66] the double tudor domain of JMJD2A[67] and the plant homeodomain (PHD) fingers of ING2,[68-71] BPTF[68,71] and Yng1.[72] Crystallographic analysis reveals that the PHD fingers of ING2,[69] BPTF[68] and Yng1[72] all contain an aromatic cage that mediates binding to methyl-lysine, a feature shared by other methyl-lysine-binding domains.[73] The structural basis of H3K4me2 or H3K4me3 binding by the PHD finger is of particularly broad significance, because this recognition domain is present in many chromatin-associated proteins that carry out histone modification.[74,75]

The ability of the PHD finger to mediate binding to H3K4me2 and H3K4me3 led several groups to examine the function of a similar domain that earlier had been identified within residues 419 through 481 of the non-core region of RAG-2.[76] This noncanonical PHD finger (Fig. 1B) was shown to mediate direct binding of RAG-2 to histone H3 di- or trimethylated at K4, with a preference for H3K4me3.[54,55] Mutations that abolish binding of the RAG-2 PHD finger to H3K4me3 (Fig. 1B) were found to impair V(D)J recombination both within extrachomosomal substrates and at endogenous loci.[54,55] Moreover, the association of the RAG-2 PHD finger with chromatin across the immunoglobulin heavy chain locus is positively correlated with the density of H3K4me3.[54] Mutations that disrupt H3K4me3 binding or Zn^{++} coordination by the RAG-2 PHD finger had been associated earlier with combined hereditary immunodeficiencies in humans,[77-81] underscoring the physiologic importance of these interactions.

The crystal structures of complexes between the RAG-2 PHD finger and modified H3 peptides have shown that this domain, while functionally related to its canonical cousins, exhibits the unusual ability to integrate epigenetic marks.[82] In the complex with a peptide bearing K4me3, the trimethyl ammonium group of K4 is buried in an "aromatic cage" similar to that of other methyl-lysine-binding domains. An important difference between the PHD finger of RAG-2 and other H3K4me3-binding domains, however, was observed: an enhanced affinity for a doubly modified histone—namely, H3 bearing both K4Me3 and a symmetrically dimethylated arginine at position 2 (R2Me2s). This is possible because the RAG-2 PHD finger lacks a side chain carboxylate that in homologous domains forms salt bridges with unmodified R2. In RAG-2 this is replaced by tyrosine, which mediates interactions with H3R2me2s.[82] An important consequence is that binding of RAG-2 to an H3 peptide bearing K4me3 is enhanced by the presence of R2Me2s.[82] While the differential affinities of RAG-2 for singly and doubly modified histone H3 could in principle contribute to locus discrimination by the V(D)J recombinase, the physiological relevance of this property remains unclear, because symmetric methylation of histone H3 R2 has as yet not been detected in vivo.

Evidence for Allosteric Regulation of V(D)J Recombinase Activity by Histone H3 Trimethylated at Lysine 4

The engagement of histone H3K4me3 by the RAG-2 PHD finger provides a bridge between one chemical mark of active chromatin and the V(D)J recombinase machinery. Paradoxically, while V(D)J recombination is profoundly impaired by a point mutation that abolishes H3K4me3 binding by the RAG-2 PHD finger, complete removal of the non-core region, including the entire PHD finger, has only a modest debilitating effect.[54,55] To reconcile these observations it has been proposed that an inhibitory domain resides within the non-core region of RAG-2 and that suppression of recombinase activity by this domain is relieved upon engagement of the PHD finger by H3K4me3 (Fig. 2). Consistent with this proposal is a crystal structure in which the RAG-2 PHD finger—in the absence of an H3K4me3 ligand—is occupied by an amino-terminal peptide encoded by the expression construct.[82] It may be that hypermethylated H3K4 does not simply act as a docking site for the recombinase but rather plays a more active role as an allosteric trigger of RAG catalysis.

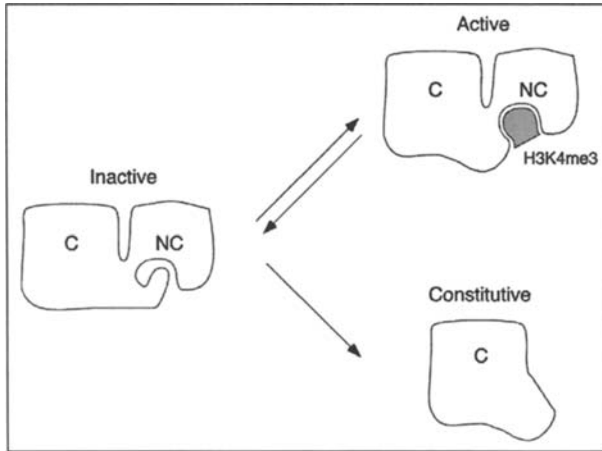


Figure 2. A model for allosteric activation of the RAG complex by modified histone H3. White figures represent RAG-2; C and NC denote core and non-core regions. respectively. Shaded object represents trimethylated lysine 4 of histone N3 (H3K4me3). In a hypothetical inactive conformation (left), the aromatic channel of the RAG-2 PHD finger is occupied by an inhibitory domain residing elsewhere in the non-core region. In the hypothetical active conformation (upper right), the PHD finger is bound by histone H3K4me3 and the putative inhibitory domain is released. The RAG-2 core fragment (lower right) lacks both the PHD finger and the putative inhibitory domain. In this configuration RAG-2 is proposed to assume an active configuration constitutively. For further discussion, see text.

Future Directions: Deposition and Integration of Epigenetic Signals Controlling V(D)J Recombination

The link between transcriptional activation and locus-specificity of V(D)J recombination has long suggested that transcription and V(D)J recombination are controlled by shared epigenetic mechanisms. Progress in understanding these mechanisms has awaited the chemical characterization of epigenetic marks and the development of methods by which the genomic distribution of these marks could be mapped. These approaches have begun to provide a detailed view of epigenetic change at antigen receptor genes as a function of development. Several important questions will continue to dominate the field.

The first is to define precisely the structural features that confer locus specificity to the V(D)J recombinase. While recognition of histone H3K4me3 by RAG-2 provides a link between active chromatin and V(D)J recombination, it is obvious that H3K4me3—a general mark of transcriptionally active chromatin—is too broadly distributed to act alone in directing the recombinase to specific sites of action. Clearly other modes of regulation must contribute to locus specificity of recombinase activity. While it seems likely that this will involve a combinatorial summation of chromatin modifications and DNA sequence elements, the answer is far from clear. A related question concerns the direct role of modified chromatin in regulating RAG activity. The proposal that the recombinase is allosterically activated upon binding of the RAG-2 PHD finger to modified chromatin will need to be tested and the relative contributions of modifications at H3K4, H3R2 and elsewhere will need to be defined. Regions of the RAG-2 other than the PHD finger may also mediate functional interactions with chromatin. The RAG-2 linker region, which lies at the amino-terminal side of the PHD finger (Fig. 1B), has been reported to bind core histones and mutations within this region were found to impair V_H-to-DJ_H joining;[83] the basis for this apparently selective effect is unclear. A third question concerns how developmental signals, such as those that emanate from the preBCR, govern deposition and removal of epigenetic marks at antigen receptor loci. A resolution of these outstanding issues will provide a starting point from which to address the larger problem of allelic exclusion.

References

1. Fugmann SD, Villey IJ, Ptaszek LM et al. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. Mol Cell 2000; 5:97-107.
2. Kim DR, Dai Y, Mundy CL et al. Mutations of acidic residues in RAG1 define the active site of the V(D) J recombinase. Genes Dev 1999; 13:3070-3080.
3. Landree MA, Wibbenmeyer JA, Roth DB. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. Genes Dev 1999; 13:3059-3069.
4. Akamatsu Y, Oettinger MA. Distinct roles of RAG1 and RAG2 in binding the V(D)J recombination signal sequences. Mol Cell Biol 1998; 18:4670-4678.
5. Difilippantonio MJ, McMahan CJ, Eastman QM et al. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. Cell 1996; 87:253-262.
6. Swanson PC, Desiderio S. V(D)J recombination signal recognition: distinct, overlapping DNA-protein contacts in complexes containing RAG1 with and without RAG2. Immunity 1998; 9:115-125.
7. Eastman QM, Villey IJ, Schatz DG. Detection of RAG protein-V(D)J recombination signal interactions near the site of DNA cleavage by UV cross-linking. Mol Cell Biol 1999; 19:3788-3797.
8. Swanson PC, Desiderio S. RAG-2 promotes heptamer occupancy by RAG-1 in the assembly of a V(D)J initiation complex. Mol Cell Biol 1999; 19:3674-3683.
9. Hiom K, Gellert M. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. Cell 1997; 88:65-72.
10. Qiu JX, Kale SB, Yarnell Schultz H et al. Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol Cell 2001; 7:77-87.
11. Cuomo CA, Oettinger MA. Analysis of regions of RAG-2 important for V(D)J recombination. Nucleic Acids Res 1994; 22:1810-1814.
12. Kirch SA, Rathbun GA, Oettinger MA. Dual role of RAG2 in V(D)J recombination: catalysis and regulation of ordered Ig gene assembly. EMBO J 1998; 17:4881-4886.
13. McMahan CJ, Difilippantonio MJ, Rao N et al. A basic motif in the N-terminal region of RAG1 enhances V(D)J recombination activity. Mol Cell Biol 1997; 17:4544-4552.
14. Sadofsky MJ, Hesse JE, Gellert M. Definition of a core region of RAG-2 that is functional in V(D)J recombination. Nucleic Acids Res 1994; 22:1805-1809.
15. Sadofsky MJ, Hesse JE, McBlane JF et al. Expression and V(D)J recombination activity of mutated RAG-1 proteins. Nucleic Acids Res 1993; 21:5644-5650.
16. Steen SB, Han JO, Mundy C et al. Roles of the "dispensable" portions of RAG-1 and RAG-2 in V(D)J recombination. Mol Cell Biol 1999; 19:3010-3017.
17. Sekiguchi JA, Whitlow S, Alt FW. Increased accumulation of hybrid V(D)J joins in cells expressing truncated versus full-length RAGs. Mol Cell 2001; 8:1383-1390.
18. Akamatsu Y, Monroe R, Dudley DD et al. Deletion of the RAG2 C terminus leads to impaired lymphoid development in mice. Proc Natl Acad Sci USA 2003; 100:1209-1214.
19. Liang HE, Hsu LY, Cado D et al. The "dispensable" portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B- and T-cell development. Immunity 2002; 17:639-651.
20. Talukder SR, Dudley DD, Alt FW et al. Increased frequency of aberrant V(D)J recombination products in core RAG-expressing mice. Nucleic Acids Res 2004; 32:4539-4549.
21. Li Z, Dordai DI, Lee J et al. A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. Immunity 1996; 5:575-589.
22. Lin WC, Desiderio S. Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. Science 1993; 260:953-959.
23. Lin WC, Desiderio S. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. Proc Natl Acad Sci USA 1994; 91:2733-2737.
24. Desiderio S, Lin WC, Li Z. The cell cycle and V(D)J recombination. Curr Top Microbiol Immunol 1996; 217:45-59.
25. Schlissel M, Constantinescu A, Morrow T et al. Double-strand signal sequence breaks in V(D)J recombination are blunt, 5'-phosphorylated, RAG-dependent and cell cycle regulated. Genes Dev 1993; 7:2520-2532.
26. Jiang H, Ross AE, Desiderio S. Cell cycle-dependent accumulation in vivo of transposition-competent complexes between recombination signal ends and full-length RAG proteins. J Biol Chem 2004; 279:8478-8486.
27. Ross AE, Vuica M, Desiderio S. Overlapping signals for protein degradation and nuclear localization define a role for intrinsic RAG-2 nuclear uptake in dividing cells. Mol Cell Biol 2003; 23:5308-5319.
28. Jiang H, Chang FC, Ross AE et al. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D) J recombinase to the cell cycle. Mol Cell 2005; 18:699-709.

29. Takata M, Sasaki MS, Sonoda E et al. Homologous recombination and nonhomologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J 1998; 17:5497-5508.
30. Feeney AJ. Factors that influence formation of B-cell repertoire. Immunol Res 2000; 21:195-202.
31. Nakase H, Takahama Y, Akamatsu Y. Effect of CpG methylation on RAG1/RAG2 reactivity: implications of direct and indirect mechanisms for controlling V(D)J cleavage. EMBO Rep 2003; 4:774-780.
32. Whitehurst CE, Chattopadhyay S, Chen J. Control of V(D)J recombinational accessibility of the D beta 1 gene segment at the TCR beta locus by a germline promoter. Immunity 1999; 10:313-322.
33. Mathieu N, Hempel WM, Spicuglia S et al. Chromatin remodeling by the T-cell receptor (TCR)-beta gene enhancer during early T-cell development: Implications for the control of TCR-beta locus recombination. J Exp Med 2000; 192:625-636.
34. Whitehurst CE, Schlissel MS, Chen J. Deletion of germline promoter PD beta 1 from the TCR beta locus causes hypermethylation that impairs D beta 1 recombination by multiple mechanisms. Immunity 2000; 13:703-714.
35. Mostoslavsky R, Singh N, Kirillov A et al. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. Genes Dev 1998; 12:1801-1811.
36. Goldmit M, Ji Y, Skok J et al. Epigenetic ontogeny of the Igk locus during B-cell development. Nat Immunol 2005; 6:198-203.
37. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85:887-897.
38. Kwon J, Imbalzano AN, Matthews A et al. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. Mol Cell 1998; 2:829-839.
39. Golding A, Chandler S, Ballestar E et al. Nucleosome structure completely inhibits in vitro cleavage by the V(D)J recombinase. EMBO J 1999; 18:3712-3723.
40. Kwon J, Morshead KB, Guyon JR et al. Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA. Mol Cell 2000; 6:1037-1048.
41. Baumann M, Mamais A, McBlane F et al. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. EMBO J 2003; 22:5197-5207.
42. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J 2001; 20:6394-6403.
43. Chowdhury D, Sen R. Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements. Immunity 2003; 18:229-241.
44. Johnson K, Calame K. Transcription factors controlling the beginning and end of B-cell differentiation. Curr Opin Genet Dev 2003; 13:522-528.
45. McMurry MT, Krangel MS. A role for histone acetylation in the developmental regulation of VDJ recombination. Science 2000; 287:495-498.
46. Tripathi R, Jackson A, Krangel MS. A change in the structure of Vbeta chromatin associated with TCR beta allelic exclusion. J Immunol 2002; 168:2316-2324.
47. Johnson K, Pflugh DL, Yu D et al. B-cell-specific loss of histone 3 lysine 9 methylation in the V(H) locus depends on Pax5. Nat Immunol 2004; 5:853-861.
48. Morshead KB, Ciccone DN, Taverna SD et al. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci USA 2003; 100:11577-11582.
49. Osipovich O, Milley R, Meade A et al. Targeted inhibition of V(D)J recombination by a histone methyltransferase. Nat Immunol 2004; 5:309-316.
50. Thomas LR, Miyashita H, Cobb RM et al. Functional analysis of histone methyltransferase g9a in B- and T-lymphocytes. J Immunol 2008; 181:485-493.
51. Sims RJ 3rd, Reinberg D. Histone H3 Lys 4 methylation: caught in a bind? Genes Dev 2006; 20:2779-2786.
52. Perkins EJ, Kee BL, Ramsden DA. Histone 3 lysine 4 methylation during the pre-B to immature B-cell transition. Nucleic Acids Res 2004; 32:1942-1947.
53. Chakraborty T, Chowdhury D, Keyes A et al. Repeat organization and epigenetic regulation of the DH-Cmu domain of the immunoglobulin heavy-chain gene locus. Mol Cell 2007; 27:842-850.
54. Liu Y, Subrahmanyam R, Chakraborty T et al. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity 2007; 27:561-571.
55. Matthews AG, Kuo AJ, Ramon-Maiques S et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature 2007; 450:1106-1110.
56. Dover J, Schneider J, Tawiah-Boateng MA et al. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J Biol Chem 2002; 277:28368-28371.

57. Sun ZW, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature 2002; 418:104-108.
58. Lee JS, Shukla A, Schneider J et al. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell 2007; 131:1084-1096.
59. Henry KW, Wyce A, Lo WS et al. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes Dev 2003; 17:2648-2663.
60. Zhang Y. Transcriptional regulation by histone ubiquitination and deubiquitination. Genes Dev 2003; 17:2733-2740.
61. Kao CF, Hillyer C, Tsukuda T et al. Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. Genes Dev 2004; 18:184-195.
62. Xiao T, Kao CF, Krogan NJ et al. Histone H2B ubiquitylation is associated with elongating RNA polymerase II. Mol Cell Biol 2005; 25:637-651.
63. Minsky N, Shema E, Field Y et al. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. Nat Cell Biol 2008; 10:483-488.
64. Kirmizis A, Santos-Rosa H, Penkett CJ et al. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. Nature 2007; 449:928-932.
65. Flanagan JF, Mi LZ, Chruszcz M et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature 2005; 438:1181-1185.
66. Sims RJ 3rd, Chen CF, Santos-Rosa H et al. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. J Biol Chem 2005; 280:41789-41792.
67. Huang Y, Fang J, Bedford MT et al. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science 2006; 312:748-751.
68. Li H, Ilin S, Wang W et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature 2006; 442:91-95.
69. Pena PV, Davrazou F, Shi X et al. Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature 2006; 442:100-103.
70. Shi X, Hong T, Walter KL et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature 2006; 442:96-99.
71. Wysocka J, Swigut T, Xiao H et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature 2006; 442:86-90.
72. Taverna SD, Ilin S, Rogers RS et al. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol Cell 2006; 24:785-796.
73. Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell 2007; 25:15-30.
74. Bienz M. The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci 2006; 31:35-40.
75. Mellor J. It takes a PHD to read the histone code. Cell 2006; 126:22-24.
76. Callebaut I, Mornon JP. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. Cell Mol Life Sci 1998; 54:880-891.
77. Gomez CA, Ptaszek LM, Villa A et al. Mutations in conserved regions of the predicted RAG2 kelch repeats block initiation of V(D)J recombination and result in primary immunodeficiencies. Mol Cell Biol 2000; 20:5653-5664.
78. Noordzij JG, de Bruin-Versteeg S, Verkaik NS et al. The immunophenotypic and immunogenotypic B-cell differentiation arrest in bone marrow of RAG-deficient SCID patients corresponds to residual recombination activities of mutated RAG proteins. Blood 2002; 100:2145-2152.
79. Schwarz K, Gauss GH, Ludwig L et al. RAG mutations in human B-cell-negative SCID. Science 1996; 274:97-99.
80. Villa A, Sobacchi C, Notarangelo LD et al. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. Blood 2001; 97:81-88.
81. Villa A, Santagata S, Bozzi F et al. Partial V(D)J recombination activity leads to Omenn syndrome. Cell 1998; 93:885-896.
82. Ramon-Maiques S, Kuo AJ, Carney D et al. The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. Proc Natl Acad Sci USA 2007; 104:18993-18998.
83. West KL, Singha NC, De Ioannes P et al. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. Immunity 2005; 23:203-212.

Chapter 14

V(D)J Recombination:
Of Mice and Sharks

Ellen Hsu*

Abstract

The adaptive immune system of jawed vertebrates is based on a vast, anticipatory repertoire of specific antigen receptors, immunoglobulins (Ig) in B-lymphocytes and T-cell receptors (TCR) in T-lymphocytes. The Ig and TCR diversity is generated by a process called V(D)J recombination, which is initiated by the RAG recombinase. Although RAG activity is very well conserved, the regulated accessibility of the antigen receptor genes to RAG has evolved with the species' organizational structure, which differs most significantly between fishes and tetrapods. V(D)J recombination was primarily characterized in developing lymphocytes of mice and human beings and is often described as an ordered, two-stage program. Studies in rabbit, chicken and shark show that this process does not have to be ordered, nor does it need to take place in two stages to generate a diverse repertoire and enable the expression of a single species of antigen receptor per cell, a restriction called allelic exclusion.

Introduction

Origins of the Adaptive Immune System

V(D)J recombination is the process by which antigen receptors, immunoglobulin (Ig) and T-cell receptor (TCR), are assembled for expression during development of the respective B- and T-lymphocytes. Somatic rearrangement of the V (variable), D (diversity) and J (joining) gene segments[1] is initiated by the recombinase RAG (recombination-activating gene)[2,3] in a cut-and-paste process that entails joining of these separate gene components to encode the V region, the N-terminus of the receptor polypeptide. The V region is 100-120 amino acid long and forms the ligand-binding site in heterodimers of heavy (H) and light (L) chains of Ig, the alpha and beta chains of TCRαβ and the gamma and delta chains of TCRγδ.

RAG and lymphocytes expressing Ig and TCR are present in all jawed vertebrates (Fig. 1), from cartilaginous fishes to mammals. Neither RAG nor the rearranging receptors are found in protochordates or lamprey and hagfish, which suggests that the present RAG function became established in a vertebrate ancestor sometime in the 80 million years between the divergence of jawless fishes and cartilaginous fishes.[4,5] Extensive duplication events, either two whole-genome duplications or one genome-wide duplication and multiple segmental duplications occurred before and after divergence of jawless fishes.[6-9] The incipience and evolution of the adaptive immune system took place during this period of extensive genomic restructuring.[10]

The origin of the rearranging genes was first suggested by Sakano and coworkers,[11] who remarked that the recognition motifs (recombination signal sequences, RSS) adjacent the V(D)J gene segments were reminiscent of signals found at the termini of integrated transposable elements.

*Ellen Hsu—Department of Physiology and Pharmacology, State University of New York, Health Science Center at Brooklyn, Brooklyn, New York, USA. Email: hsue@hscbklyn.edu

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

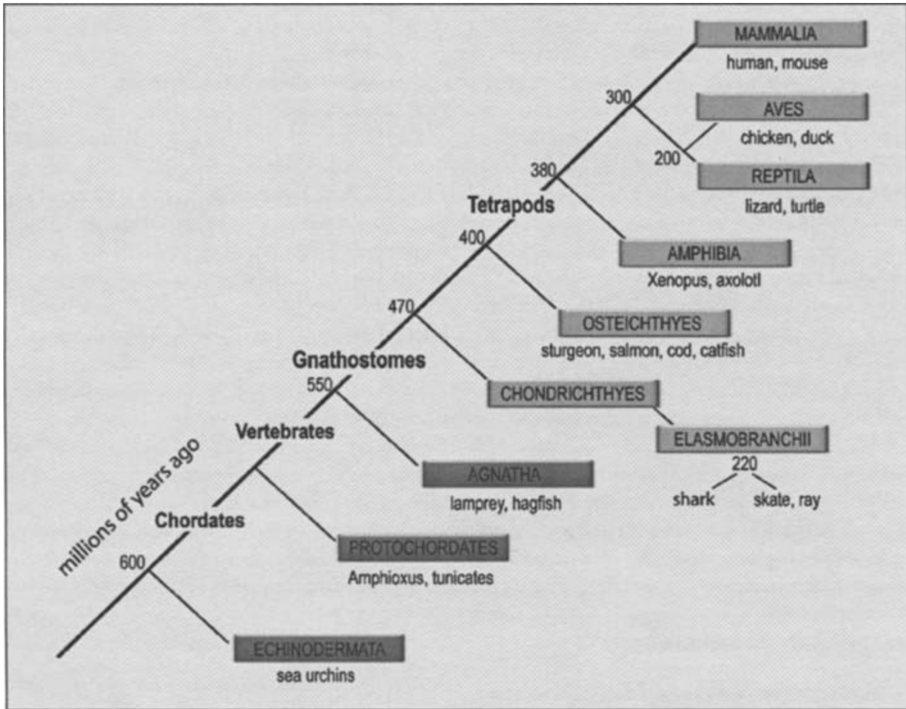


Figure 1. Evolution of the chordates. The phylogenetic relationships among chordates is shown (boxes) with notations of the major animal models in each taxon beneath the boxes. The adaptive immune system, defined by RAG-mediated rearranging antigen receptor genes of the Ig superfamily and by the major histocompatibility complex, has been found only in the jawed vertebrates (gnathostomes, beige boxes). Protochordates include Cephalochordates (Amphioxus) and Urochordates (Ciona/tunicates). Numbers denote when the taxa emerged in evolution (millions of years ago). Reprinted with some alterations; Hsu E, Pulham N, Rumfelt LL et al. The plasticity of immunoglobulin gene systems in evolution. Immunol Rev 2006; 210:8-26. Copyright Blackwell Munksgaard 2006.

Subsequently the chemistry of the RAG-mediated pathway was found to resemble those described for transpositional recombination by mobile elements.[12] With the growing availability of genome information from different species, it became feasible to attempt delving into the origins of RAG and its recognition sequences. Terminal inverted repeats with motifs and spacer interval similar to RSS were observed in the *Transib* transposon in nematodes, insects and sea urchin.[13] RAG is composed of two components and both RAG1- and RAG2-like sequences were detected in the sea urchin genome, although their function is not yet clear.[14] These discoveries, together with demonstration of latent transposase activity in RAG,[15,16] argue for RAG having been part of a DNA transposon that was introduced early into the vertebrate lineage, evolving to its role of V(D)J recombinase by retaining the excision component of transposase activity.[17-19] The presence of RAG sequences in echinoderms could indicate entry of the transposon at a far earlier time and lost in certain phyla and classes (protochordates, jawless fishes) but retained in others, or else a separate horizontal transfer in jawed vertebrates.

It is hypothesized that in the ancestral vertebrate the RAG transposon became integrated into a V-like gene, splitting it into two components that can rejoin after RAG-induced double-strand breakage and removal of the intervening DNA.[11] The cleavage occurs in the same place due to RSS recognition, but because of the nucleotide loss and/or gain arising from the repair process, the new

joints would be varied in sequence (next section). Breakage and repair of DNA induced by RAG thus generates molecular heterogeneity, which may arguably have been the selecting factor if the original V gene had an immune function that was enhanced by diversified sequences.

V(D)J recombination became established in early vertebrates about 500 million years ago and is the process that assembles Ig and TCR genes in all species and the species-specific receptor genes like the IgNAR (new antigen receptor) and NAR-TCR in sharks[20,21] and TCRμ in marsupials.[22] This chapter deals mainly with comparative studies on the Ig gene system. There is considerably more information on antibody in early vertebrates, due to the much longer history of studies of Ig protein and to the relative ease of detecting VH sequences across species with heterologous probes. The IgM molecule is very well conserved from sharks to mammals[23,24] in overall sequence and structure, being the antigen receptor on naive B-cells that in plasma cells is secreted as a polymeric antibody, usually a pentamer. TCR cDNA sequences characterized in all animals show that they are cell surface receptors only. TCRαβ and TCRγδ have been cloned from all classes of jawed vertebrates,[25-29] including all three major groups of mammals (marsupials such as opossum,[30] monotremes such as duckbill platypus[31] and placentals of most orders, including rodents, rabbits, ruminants and primates).

There are two evolutionarily conserved features of V(D)J recombination: the mechanism of RAG action and the regulation of this process to ensure one end result—that only one kind of antigen receptor is expressed per cell (for a review, see ref. 32). This restriction is called allelic exclusion. Although the recombination pathway mediated by RAG is well conserved, the regulated accessibility of the antigen receptor genes to RAG has evolved with the organizational structure, which differs most significantly between cartilaginous fishes and tetrapods (Fig. 2).

V(D)J Rearrangement

RAG Recognition and Joint Resolution

The rearranging elements—the gene segments V, D and J with their adjacent RSS—are present in all classes of jawed vertebrates, as are the key enzymes involved in DNA nicking and modification, RAG1/RAG2 and terminal deoxynucleotidyl transferase (TdT). Although these lymphocyte-specific enzymes have been studied almost entirely in mouse or in vitro systems,[33] their highly conserved mode of action in other animals may be deduced. First, pairwise recognition of the RSS is required and the RSS pair to be recombined must consist of one RSS containing a 12-bp spacer and the other a 23-bp spacer ("12/23 rule").[1] In all vertebrates where the genomic organization of the gene segments has been determined, the RSS that flank potentially recombinogenic gene segments reflect this pairing relationship.

RAG initiates the pathway that leads to double-strand breaks at either gene segment and the coding ends being subsequently joined by the cell's DNA repair processes (Fig. 3). Double-strand breakage is obtained in a transesterification reaction that results in a covalently closed hairpin on the coding end and a free blunt RSS at the other. The hairpin is opened asymmetrically, creating an overhang with inverted repeat, some of which is occasionally retained (P region) as part of the ligated joint.

The presence of P region is thus indicative of a hairpin intermediate created during the double-strand break and joining process. Examination of the VD and DJ junctions in Ig H chains and TCR β and δ chains, or VJ junctions in Ig L chains and TCR α and γ chains, the portion of the V sequence called CDR3 (complementarity-determining region 3), shows germline contribution (V and J gene segment flanks, portions of D gene sequence) and occasional P region in all animals, suggesting that V(D)J recombination at different loci and in various species undergo the same unique process involving hairpinned coding ends.

Selection for Junctional Diversification

A second category of somatically-generated additions at the junction is N region, which consists of nontemplated, mostly GC-rich sequences catalyzed by TdT[34,35] that, together with coding end-processing mediated by exonucleases, are the main contributors to generating the diversification at the junctions of the rejoined gene segments.

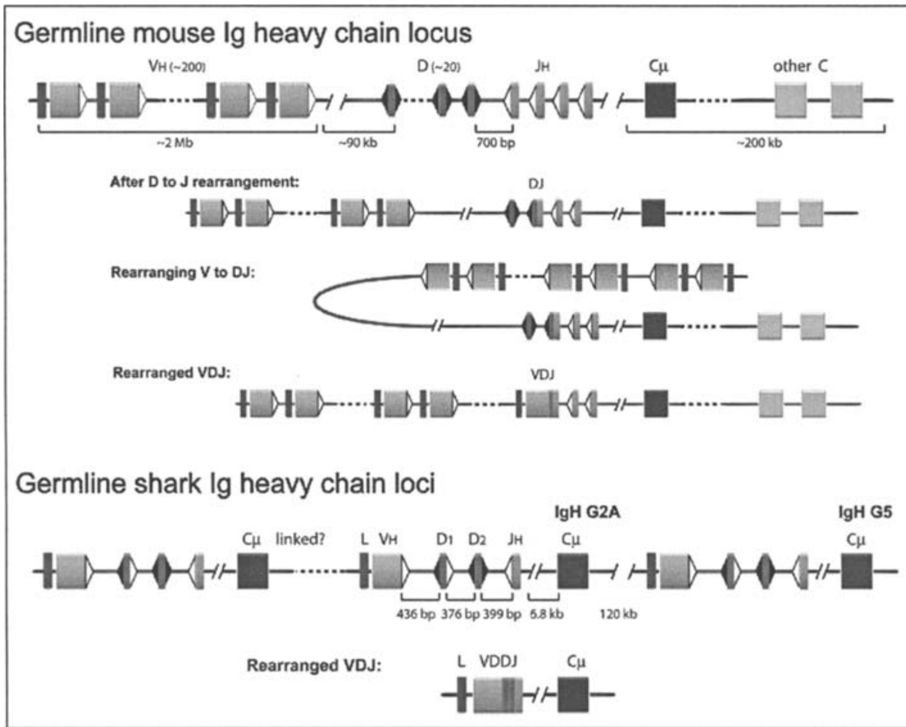


Figure 2. Comparison of Ig H chain genes in mouse and shark. Germline mouse Ig H chain locus: the mammalian H chain locus consists of a series of tandemly duplicated V_H, D and J_H gene segments that rearrange during B-cell development. The recombined VDJ is transcribed with one of the downstream constant (C) region genes, here simplified as single units (blue box is Cμ). The V_H is represented by olive boxes, preceded by the leader sequence in dark green and flanked by the recombination signal sequence (RSS, white triangle) at the 3′ end, that consists of heptamer and nonamer motifs separated by a 23 bp spacer sequence. As indicated, the distance between the 3′-most V_H and the first functional D is 90 kb. The D gene segments in red, flanked on both sides by RSS (black triangles) containing 12 bp spacers and the J_H gene segments (orange) with 23 bp spacer RSS. After D to J rearrangement: the first stage of rearrangement involves recombination between D and J_H, with the intervening DNA excised. The DJ product is depicted as a fusion of the red and orange boxes, with the RSS flanking its 5′ end. Rearranging V to DJ: locus contraction and looping of the DNA allows linearly distant V_H gene segments to recombine with the DJ. The final VDJ product is shown as Rearranged VDJ. Germline shark Ig H chain loci: the IgM H chain genes in sharks and skates (cartilaginous fishes) are multiple miniloci each consisting of V_H, two D, one J_H and one Cμ gene (blue box). The gene segments in any nurse shark IgH gene are located about 400 bp apart as shown but are distant (e.g., 6.3-6.8 kb) from the Cμ1 exon. The physical relationships among the loci are not clear except for one instance, where they were located 120 kb apart.[72] Rearranged VDJ: the four gene segments rearrange within the minilocus to VDDJ (called VDJ). Whereas in mouse IgH gene rearrangement takes place in a strict order (D to J_H before V_H to DJ), the rearrangement of the four gene segments in the shark takes place at once and without any strict order. Reprinted with permission from Malecek K, Lee V, Feng W et al. Immunoglobulin heavy chain exclusion in the shark. PLoS Biol 2008; 6:e157. Copyright 2008 Malecek et al. A color version of this image is available at www.landesbioscience.com/curie.

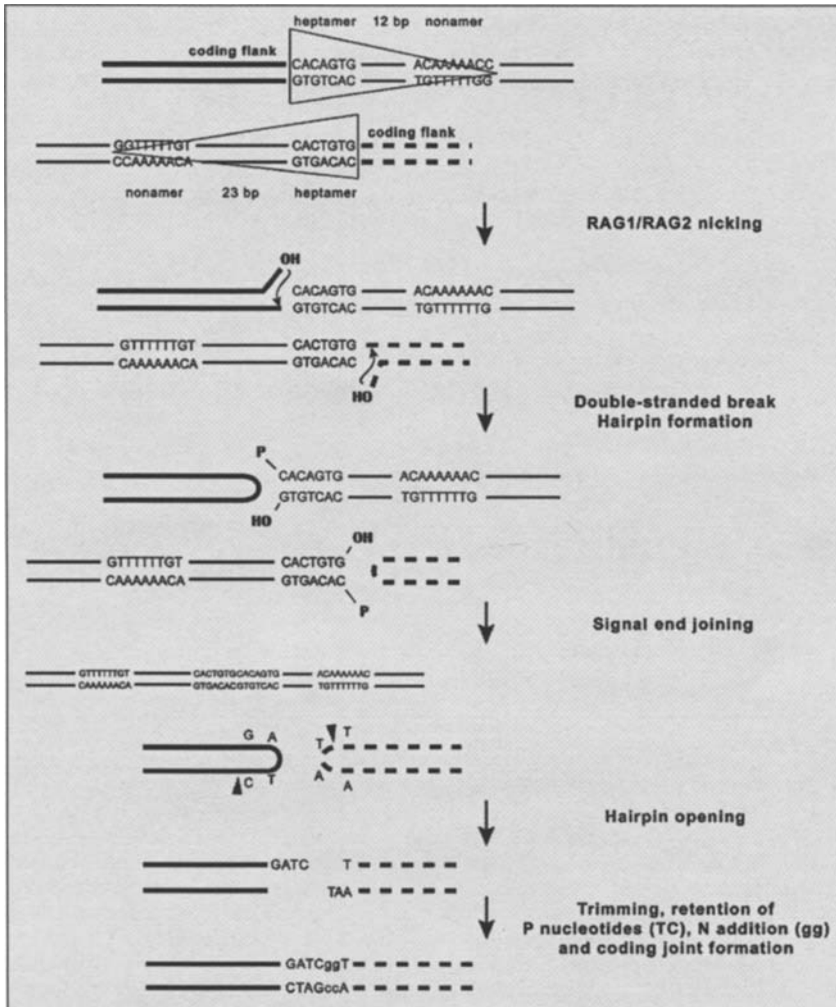


Figure 3. RAG-mediated recombination. Details of V(D)J recombination are described in other chapters of this book. The flanks of gene segments before rearrangement are shown with the RSS enclosed by triangles, to correlate with the symbols in Figures 2 and 4. The RSS pair is bound by RAG, which introduces nicks. The nicking occurs 5′ of the 7-mer end of the RSS on the top strand of each of the two Ig gene segments, producing a 3′-hydroxy on the coding end of the Ig gene segment and a 5′ phosphoryl on the RSS (signal) end. The result is a duplex nicked at either RSS. The second step involves intramolecular transesterfication reactions where the 3′-OH attack the opposing phosphodiester bonds, causing the coding ends to become a covalently closed hairpins and freeing the blunt signal ends. Joining of the ends is carried out by the non-homologous end joining repair pathway. The hairpin coding ends are opened asymmetrically by the nuclease Artemis and the resultant single-stranded overhang consists of a portion of the coding end and its complementary sequence. Sometimes the overhang could be included as part of the final joined product and is observed as inverted repeat sequence (P region). The DNA ends are trimmed; TdT may insert nontemplated nucleotides (lower case letters). Reprinted with some alterations; Hsu E. Immunoglobulin recombination signal sequences: somatic and evolutionary functions. In: Caporale L, ed. The Implicit Genome, New York: Oxford University Press, 2005, Chapter 9. Copyright 2006 by Oxford University Press, Inc.

The other members of the mammalian Family X DNA polymerases, to which TdT belongs, are not restricted to precursor lymphocytes, but two of them, Pol μ and Pol λ, are also involved at different points during coding end-processing and appear to modulate the extent of coding end nucleolytic processing.[36] TdT and Pol μ are very closely related[37] and their presence in fishes,[38] in contrast to the single-copy ancestral form in the urochordate Ciona (tunicates, Fig. 1), suggests that the lymphocyte-specific TdT evolved to its current role in the immune system by the time of divergence of cartilaginous fishes. The early involvement of TdT in the evolution of V(D)J recombination reflects its importance in amplifying the selected attributes conferred by gene rearrangement: sequence and sequence length variation as a result of RAG-induced breakage.

Diversification mechanisms like mutation or gene conversion may exist in invertebrates[39] and predate the rearranging Ig gene system, but these processes do not generate sequence length diversity repeatedly and reliably in one location that will tolerate a loop size spectrum of 2-23 amino acids in the human H chain[40] or 9-13 in the shark L chain.[41] The greatest contribution to the combining site topology is thus made by the variable CDR3. From crystallographic studies of antigen-antibody complexes H chain CDR3 appears to play the most significant role, not only in the number of contacts with antigen but also in its potential for conformational changes for "induced fit" upon binding ligand.[42,43]

Novel Rearranging Genes in Sharks and Marsupials

Some species require antigen receptor diversity additional to that provided by heterodimer specificities of the Ig/TCR repertoire. They are (1) shark IgNAR[20]: neither Ig nor TCR but an early divergent gene, (2) shark NAR-TCR[21]: TCR isoforms, produced by grafting an additional V region onto an existing TCRδ rearrangement by splicing, forming two successive V regions and (3) marsupial TCRμ[22]: a hybrid of Ig and TCR components whose product also contains two joined V regions. IgNAR is a secreted serum protein and the other two presumably are active in cell-mediated processes.

IgNAR is a dimer but the V regions are not paired; the ligand-binding site is thus a single V region. The IgNAR V region is generated by four rearrangements—V, three D and J gene segments—providing highly variable and exceptionally plastic CDR3 that are postulated to adopt multiple conformations for induced-fit binding.[44] Ig H chain dimers with single-domain V regions (VHH) are also expressed in camels,[44] although these gene segments are part of the IgH locus.[45] Shark IgNAR and the camel VHH are the result of convergent evolution, as are shark NAR-TCR and opossum TCRμ. NAR-TCR is part of the shark TCRδ locus, whereas TCRμ in opossum are encoded by independent gene clusters.[46]

TCRμ, like IgNAR, involves rearrangement of 2-3 D elements. Because TCRμ and IgNAR are encoded by a few miniloci, their repertoire is based solely on CDR3 junctional diversity. The use of a single V domain, some with longer CDR3, in cartilaginous fishes and in mammals suggests that there exists some category of antigens that require ligand-binding sites perhaps more flexible than provided by the classical Ig or TCR heterodimer.

One Receptor Per Cell

Both the strength and the weakness of V(D)J recombination is its random nature. An immensely diverse, anticipatory repertoire is generated concomitant with cell and resource wastage. The process cannot ensure that the V becomes joined in-frame with respect to the J (and C region) sequence, so that least two of three rearrangements attempts are nonfunctional. Moreover, randomly-generated specificities also include those that recognize self components and these are eliminated at the immature lymphocyte stage when triggered by a self ligand. Selection for self-tolerance or mounting an immune response is most efficaciously (i.e., specifically) mediated when only one species of receptor is expressed per cell.[47] The last phenomenon, known generally as allelic exclusion, results from regulated RAG access to the recombinogenic elements. V(D)J recombination is lineage- and cell stage-specific, meaning that DNA from nonlymphoid cells or from cells of the incorrect developmental stage are not acted upon by RAG.[48] There are, however, some interesting exceptions in cartilaginous fishes and these are described in a later section ("Rearrangement of Ig genes in non-B-cells").

V(D)J Rearrangement Patterns

In the mouse and human systems the rearrangement of Ig H and L chains is often described as an ordered, regulated program.[49] The well-studied steps involve formation of the DJ before recombination of VH to the DJ in pro B-cells, followed by cell division and subsequent rearrangement of the L chain genes in pre B-cells, where the kappa L chain (Igκ) locus is activated before the second L chain isotype, lambda (Igλ). The regulated accessibility of different genes and gene segments to RAG enable one H chain allele to be expressed (allelic exclusion, H chain exclusion) with one allele of either κ or λ L chain (allelic and isotypic exclusion); hence, one kind of antigen receptor per lymphocyte. Outside of the mouse model there is currently little information on TCR or Ig chromatin and DNA modification, but V(D)J rearrangement patterns reflect the order of gene accessibility to RAG and these are compared among mouse, rabbit, chicken and shark.

Mouse

In the mouse IgH gene rearrangement takes place in a set order and in step-wise fashion.[59,51] At the pro B-cell stage the chromatin domain encompassing the D, JH and Cμ genes become activated, probably through the intronic enhancer and allow D to JH recombination on both chromosomes. This is followed by activation of the chromatin domain containing the upstream VH genes. Because of the very large distance between the VH gene segments and the DJ, locus contraction and looping of the DNA[52-54] are required to bring them into close proximity for rearrangement (Fig. 2). H chain exclusion is the outcome of the staggering of the V to DJ step between the two alleles. If the first VDJ is not viable, rearrangement continues on the homologous chromosome.

The initiation and maintenance of allelic exclusion involves relocation of the genes in nuclear compartments.[53,55] In pro B-cells IgH repositions away from the nuclear periphery and this may have to do with its activation; in pre B-cells the nonrearranged allele is recruited to the pericentrometric heterochromatin, an interaction thought to be repressive for recombination. How rearrangement begins at one allele before the other is not clear and the basis may differ at the IgH, TCRβ and Igκ genes. An explanation for asynchronous rearrangement at the TCRβ locus has been recently proposed after finding that in rearranging T-cells both alleles of TCRβ interacted with repressive nuclear compartments at equal and high frequency.[56] This observation suggests a limited window of opportunity to achieve the V to DJ step and that any rearrangement is consequently a very low frequency event. Two simultaneous rearrangements in a cell are thus unlikely to occur and allelic inclusion is avoided.

Rabbit

Mouse IgH configurations in hybridoma cell lines[57] reflect the frequency of the recombination events, which render 51% of them VDJ/DJ, 44% VDJ/VDJ- and 5% VDJ/germline (VDJ is the expressed rearrangement, VDJ- is nonfunctional). In contrast, the IgH configurations in rabbit cell lines were: 40% VDJ/DJ, 10% VDJ/VDJ- and 50% VDJ/germline.[58] Since the D to JH step occurred on both alleles in 95% of mouse B-cells, the finding that it has not done so in 50% of rabbit B-cells suggests that asynchrony between IgH alleles can exist to a greater extent in rabbit. Lanning and coworkers[59] hypothesized that the D to JH rearrangement in rabbit involves slower kinetics and is the rate-determining step; once DJ is achieved on one chromosome there is rapid recombination to VDJ. Because of the overall inefficiency of the D to JH step, the relatively few numbers of cells with VDJ/VDJ- reflect a restricted time opportunity for the laggard allele to achieve VDJ.

Although some infrequent VH to D rearrangement was observed in rabbit splenocytes,[58] its significance is unknown, since the recombined DJ is the primary intermediate isolated from pro-B-cells. Cloned fetal rearrangements carried the two D-proximal VH genes, VH1 and the neighboring pseudogene VH2,[60] despite >100 available VH upstream; this early rearrangement bias together with clonal expansion of B-cells with VH1-expressing VDJ causes such H chains to be 70-90% of expressed Ig molecules.[61] Usage of the D-proximal VH1 in rabbit can be likened to the preferential rearrangement of the D-proximal VH genes in fetal mouse liver,[62,63] but the molecular basis of either remains be to elucidated.

Chicken

The earliest recombined cells are in the yolk sac at day 5 and 6 of incubation and carry DJ only; VDJ is found on day 9.[64] Rearrangement occurs exclusively to the D-proximal VH gene, the only functional gene out of multiple VH elements; the other VH act as donor templates during the gene conversion process in the bursa. The primary Ig repertoire in chicken, as in rabbits, is generated by postrearrangement gene conversion.[65,62] There is a distinct DJ/DJ step that is B-lineage specific in chicken and this is followed by simultaneous V rearrangement at the H and L chain loci.[64,66] Clones carrying only the VJ or only the VDJ could be observed,[66] showing that there is no ordered H and L chain rearrangement, as there exists in mouse and rabbit.[61] Thus, in the chicken, L chain rearrangement is not dependent on the success of H chain rearrangement and there is no pre-B-cell stage as in mammals.

More than 90% of bursal follicles contain the VDJ/DJ configuration and none carried VDJ rearrangements on both chromosomes. Similarly, only one allele of the L chain recombined. It was suggested that the V rearrangement occurs after removal of repression from one allele randomly and that this is an event of such low efficiency that there is little probability of its occurrence on both alleles.[67]

Multiple IgH Loci in Other Vertebrate Species

The contrasting examples of mouse, rabbit and chicken show that the V(D)J recombination program is adapted for each species. There is at least one step that is limited by RAG accessibility and/or time constraints[67] and the factors that determine these parameters remain to be elucidated. These three systems all involve a choice of two H chain alleles, but when one recombination step tends to be limiting or occurring at very low frequency, then the presence of additional alleles—one or two, equally subjected to the constraints—would not greatly increase the chances for allelic inclusion. Model systems genetically manipulated to carry multiple H chain genes (interspecies hybrid tetraploid and triploid Xenopus[68] and mice triallelic for IgH[69]) do exhibit monoallelic H chain expression and thus the same would be expected for those animals with more than one naturally-occurring IgH locus. Polyploid Xenopus species carry multiple active IgH genes.[70] Bony fish, alone of all vertebrate classes, underwent an additional genome-wide duplication[71] and some species support more than one IgH locus although in most only one remains.

Ig Rearrangement in the Shark

The IgH minilocus organization in cartilaginous fishes, representatives of the earliest vertebrates, is considered primitive and ancestral to the classical IgH locus in other vertebrates. Sharks, rays and skates carry 15-200 miniloci ("clusters") each consisting of a few gene segments (VH-D1-D2-JH-Cμ)[4,23] as shown in Figure 2. In most species the rearranging elements are located within a total span of 2 kb. The clusters themselves are located far apart from each other,[72] >120 kb and can be situated on different chromosomes.[73] V(D)J recombination takes place among the four gene segments of the minilocus; there is no evidence for intercluster rearrangement in B-cells and hence no need for locus contraction in such a system. The close proximity of the gene segments (400 bp apart) also makes unlikely any separately activated chromatin domains within a cluster. In fact, there is no strict order of rearrangement of the VH, D1, D2 and JH. Once an IgH gene is activated in a precursor B-cell, its gene segments recombine all at once and to completion.[74]

In single B-cell studies, few Ig transcripts[75] and few genomic rearrangements[74] were observed per lymphocyte. In the nurse shark there are 9-12 functional IgH genes and in any B-cell there are 1-3 VDJ genomic rearrangements of which only one appeared to encode a viable receptor. Less than 10% of the cells carried any partially rearranged genes and the rest of the IgH genes were in germline configuration. This suggests that once initiated, recombination occurs efficiently between the four gene segments. These data show that H chain exclusion exists in the shark, despite its unique IgH organization. As in higher vertebrates, H chain exclusion in sharks is based on limitation of rearrangement, but the mechanism of repression (or activation) must accommodate the large and varied numbers of IgH loci in different cartilaginous fish species.

The process producing monoallelic Ig H chain expression at the murine IgH locus evolved with and is a consequence of the complex gene organization, whose multiple gene segments are scattered over 2 Mb. If you take away the locus contraction and the separately activated domains, what shark and mouse have in common is that initiation of rearrangement is an inefficient, low frequency event. Whether there are regulatory features in common between shark and tetrapod IgH gene systems remains to be established. However, a few conclusions can be extracted. Because of the large number of IgH loci and their dispersed locations, it is unlikely that H chain exclusion in the shark is based on any mechanism that predetermines[77,78] rearrangement preference at homologous chromosomes. In nurse shark at least two IgH genes are adjacent[72] and the model for kappa L chain exclusion based on rearrangement preference evinced by the earlier replicating chromosome will not distinguish multiple, linked genes. It is not clear whether the 1-3 rearrangements in a B-cell occurred simultaneously or sequentially and we suggest that their activation was probably stochastic. If it happened that one rearrangement at a shark IgH gene is nonfunctional it seems unlikely that its allele is more apt to be the one next (or simultaneously) targeted for recombination than an adjacent or any other IgH in the genome.

Rearrangement of Ig Genes in Non-B-cells

There exist pre-rearranged Ig genes in the germline of cartilaginous fishes, catfish and chicken.[78-80] In sharks, skates and rays some IgH clusters carry partially or fully recombined VD-J or VDJ and the IgL clusters joined VJ.[4] Examination of nurse shark L chain junctions in some germline-joined VJ showed P region sequence that may indicate a one-time hairpin formation. This evidence and the fact that the "12/23 rule" is always obeyed, suggest that there was RAG activity in germ cells of some animals.[81,19] It was hypothesized that RAG-mediated changes in germline Ig genes produced the VD templates used in chicken H chain gene conversion or perhaps generated D elements during antigen receptor gene evolution.

The function of recombined genes in the shark antibody repertoire is not known; it appears that many are pseudogenes. In a species with many pre-rearranged VDJ there would be a strong likelihood for allelic inclusion if more than one IgH is activated at a time, but at the moment the germline genes in these animals have not been fully characterized. Nurse shark is an instance where all its IgM clusters have been characterized and none are pre-rearranged, showing that germline-joined genes are particular to the species.[72]

Once initiated, somatic rearrangement in B-cells leads to VDJ. Partially rearranged IgH on the other hand have been observed in abundance in nurse shark thymocytes and 3-7 can be isolated per cell.[74] Thymic H chain transcripts could not be detected, implying that availability of DNA to RAG does not require transcription. That many thymocyte rearrangements are incomplete as VD-D-J, V-DDJ, etc., suggests that transcription may be part of the process that recruits[82] RAG to its target for efficient recombination. This IgH rearrangement-permissive state in thymocytes may have characteristics in common with that in germ cells enabling RAG, when present, to effect recombination. However, the state of the IgH chromatin in either cell type has yet to be characterized.

About L Chain

In the course of evolution, whole-locus duplications produced the multiple cluster organization of cartilaginous fish IgH and IgL, whereas successive tandem duplications of the gene segments V, (D) and J generated the "translocon" organization that exists in tetrapods. While H chain genes are organized either as translocon or multiple clusters, the evolution of L chain genes[83] is more complex.

The number of L chain isotypes varies among vertebrates. In chicken there is only the one locus, Igλ; in mammals there are two, Igλ and Igκ. In Xenopus there are three: Igσ (sigma) and the homologs of Igκ (Igρ, called rho) and Igλ (called Type III). In shark there are four: cartilaginous fish-specific "Igσ-cart" (called Type I/NS5) and the homologs of sigma, Igκ (called Type III/NS4) and Igλ (called Type II/NS3). Igκ is thus present in all animals except birds and its organization varies considerably. In tetrapods Igκ is one locus. In nurse shark the Igκ homolog exists as >60

miniloci, with one V, one J and one C exon and tend to be separated by some distance. However in a bony fish like zebrafish, Igκ genes (Type 1/3)[84] are arranged closely in serial arrays (examples in Fig. 4) and on at least four different chromosomes.[85]

It is not clear how L chain expression is regulated in zebrafish (or any bony fish). In cod it was shown that multiple enhancers existed in the serial clusters but not every IgL C region was associated with downstream enhancer activity.[86] It cannot be anticipated from mere distance how regulatory control is exercised. Because there can be additional possibilities for intracluster rearrangement following an initial V to J attempt (Fig. 4), we have suggested that the bony fish organization allows for correction not only of nonproductive VJ but also in-frame VJ that contribute to forming a self-reactive specificity.[84] In other words, there exists a potential for receptor editing[87] in fishes, since the organizational set-up appears to allow for secondary rearrangements.

In zebrafish, the IgH organization is translocon like tetrapods[88] so that both types of arrangement exist for its Ig genes. It is clear that H and L chain gene organizations do not have to co-evolve—as they did not in bony fish[89]—and information from this and the other model systems suggest they can be regulated independently. L chain exclusion is not as stringent as H

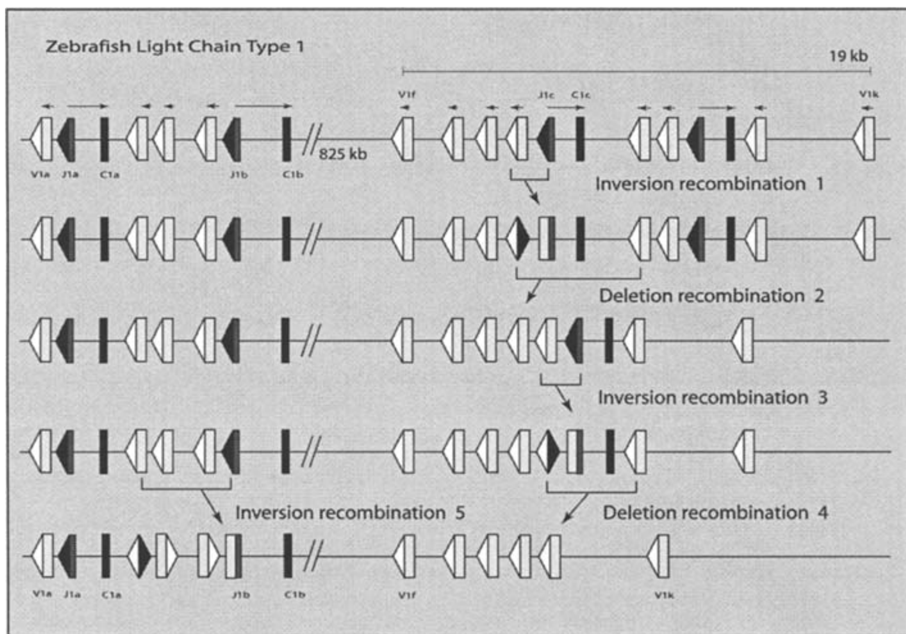


Figure 4. Organization of representative genes encoding zebrafish L chains. Some L chain Type 1 clusters on chromosome 24 are represented on top line; the names of segments are some of those identified in reference 84; their updated linkage, polarity and distances were obtained from the Zv7 zebrafish genome assembly (www.ensembl.org) and reference 85. V (yellow boxes) and J (blue) gene segments are flanked by RSS (white triangle is RSS with 12 bp spacer, black triangle is RSS with 23 bp spacer) and C exons are depicted by black boxes. The transcriptional polarities are indicated by overhead arrows. A hypothetical series of rearrangements is depicted. Inversion recombination 1: rearrangement between J1c and V1i immediately upstream to form VJ (indicated as fused rectangles) and blunt-end joined RSS (fused triangles). Deletion recombination 2: The RSS-23 of the fused signal joint recombines with downstream V gene segment and deletes intervening DNA. Inversion recombination 3: the remaining J rearranges to upstream V, forming again VJ and blunt-end joined RSS. This VJ can be excised by deletion recombination 4 and replaced by rearrangement at another cluster, inversion recombination 5. A color version of this image is available at www.landesbioscience.com/curie

chain, but the mechanism for restricting their expression in zebrafish must manage a large array of clusters, many of which carry multiple recombinogenic elements on either side of the C exon. How V(D)J recombination is sorted out in zebrafish will elucidate those aspects of RAG accessibility that evolve with individual species' immune system requirements.

Conclusion

V(D)J rearrangement was established in an ancestral jawed vertebrate about 500 million years ago. From sharks to mammals two features are evolutionarily conserved—the mechanism of RAG recombinase action and a process for limiting rearrangement activity in order to produce monospecific lymphocytes. The regulated accessibility of antigen receptor genes to RAG was characterized in precursor lymphocytes of mice and human beings, where it is usually described as an ordered, two-stage program. However, a comparison of Ig rearrangement patterns from rabbit, chicken and shark shows that this process neither has to be strictly ordered nor must take place in two stages to generate a diverse repertoire and bring about allelic exclusion.

Acknowledgements

I wish to thank Louis Du Pasquier and Martin Flajnik for reading the manuscript and Karolina Malecek for her help with the figures. This work was supported by grants from the National Institutes of Health and the National Science Foundation.

References

1. Tonegawa S. Somatic generation of antibody diversity. Nature 1983; 302:575-581.
2. Schatz DG, Oettinger MA, Baltimore D. The V(D)J recombination activating gene, RAG-1. Cell 1989; 59:1035-1048.
3. Oettinger MA, Schatz DG, Gorka C et al. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 1990; 248:1517-1523.
4. Rast JP, Litman GW. Towards understanding the evolutionary origins and early diversification of rearranging antigen receptors. Immunol Rev 1998; 166:79-86.
5. Flajnik MF, Du Pasquier L. Evolution of innate and adaptive immunity: can we draw a line? Trends Immunol 2004; 25:640-644.
6. Spring J. Genome duplication strikes back. Nature Genetics 2002; 31:128-129.
7. Furlong RF, Holland PWH. Were vertebrates octoploid? Philos Trans R Soc Lond B Biol Sci 2002; 357:531-544.
8. Holland PWH, Garcia-Fernandez J, Williams NA et al. Gene duplications and the origins of vertebrate development. Development Suppl 1994; 125-133.
9. Escriva H, Manzon L, Youson J et al. Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. Mol biol Evol 2002; 19:1440-14506.
10. Kasahara M, Suzuki T, Du Pasquier L. On the origins of the adaptive immune system: novel insights from invertebrates and cold-blooded vertebrates. Trends Immunol 2004; 25:105-111.
11. Sakano H, Huppi K, Heinrich G et al. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. Nature 1979; 280:288-294.
12. van Gent DC, Mizuuchi K, Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration. Science 1996; 271:1592-1594.
13. Kapitonov VV, Jurka J. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. PLoS Biol 2005; 3:e181.
14. Fugmann SD, Messier C, Novack LA et al. An ancient evolutionary origin of the Ragl/2 gene locus. Proc Natl Acad Sci USA 2006; 103:3728-3733.
15. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature 1998; 394:744-751.
16. Hiom K, Melek, M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell 1998; 94: 463-470.
17. Thompson CB. New insights into V(D)J recombination and its role in the evolution of the immune system. Immunity 1995; 3:531-539.
18. Fugmann SD, Lee AI, Shockett PE et al. The RAG proteins and V(D)J recombination: complexes, ends and transposition. Annu Rev Immunol 2000; 18:495-527.
19. Lewis SM, Wu GE. The old and the restless. J Exp Med 2000; 191:1631-1636.
20. Greenberg AS, Avila D, Hughes M et al. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. Nature 1995; 374:168-173.

21. Criscitiello MF, Saltis M, Flajnik MF. An evolutionarily mobile antigen receptor variable region gene: doubly rearranging NAR-TcR genes in sharks. Proc Natl Acad Sci USA 2006; 103:5036-5041.
22. Parra ZE, Baker ML, Schwarz RS et al. A unique T-cell receptor discovered in marsupials. Proc Natl Acad Sci USA 2007; 104:9776-9781.
23. Flajnik MF. Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol 2002; 2:688-698.
24. Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu Rev Immunol 1999; 17:109-147.
25. Rast JP, Anderson MK, Strong SJ et al. α, β, γ and δ T-cell antigen receptor genes arose early in vertebrate phylogeny. Immunity 1997; 6:1-11.
26. Wilson MR, Zhou H, Bengtén E et al. T-cell receptors in channel catfish: structure and expression of TCR alpha and beta genes. Mol Immunol 1998; 35:545-557.
27. Haire RN, Kitzan Haindfield MK et al. Structure and diversity of T-lymphocyte antigen receptors alpha and gamma in Xenopus. Immunogenetics 2002; 54:431-438.
28. André S, Kerfourn F, Affaticati P et al. Highly restricted diversity of TCR delta chains of the amphibian Mexican axolotl (Ambystoma mexicanum) in peripheral tissues. Eur J Immunol 2007; 37:1621-1633.
29. Kubota T, Wang J, Göbel TW et al. Characterization of an avian (Gallus gallus domesticus) TCR alpha delta gene locus. J Immunol 1999; 163:3858-3866.
30. Parra ZE, Baker ML, Hathaway J et al. Comparative genomic analysis and evolution of the T-cell receptor loci in the opossum Monodelphis domestica. BMC Genomics 2008; 9:111.
31. Parra ZE, Arnold T, Nowak MA et al. TCR gamma chain diversity in the spleen of the duckbill platypus (Ornithorhynchus anatinus). Dev Comp Immunol 2006; 30:699-710.
32. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. Nat Immunol 2003; 4:624-630.
33. Gellert M. V(D)J recombination: RAG proteins, repair factors and regulation. Annu Rev Biochemistry 2002; 71:101-132.
34. Komori T, Okada A, Stewart V et al. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. Science 1993; 261:1171-1175.
35. Gilfillan S, Dierich A, Lemeur M et al. Mice lacking TdT: mature animals with an immature lymphocyte repertoire. Science 1993; 261:1175-1178.
36. Bertocci B, De Smet A, Weill JC et al. Nonoverlapping functions of DNA polymerases mu, lambda and terminal deoxynucleotidyltransferase during immunoglobulin V(D)J recombination in vivo. Immunity 2006; 25:31-41.
37. Domínguez O, Ruiz JF, Laín de Lera T et al. DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. EMBO J 2000; 19:1731-1742.
38. Beetz S, Diekhoff D, Steiner LA. Characterization of terminal deoxynucleotidyl transferase and polymerase mu in zebrafish. Immunogenetics 2007; 59:735-744.
39. Zhang SM, Adema CM, Kepler TB et al. Diversification of Ig superfamily genes in an invertebrate. Science 2004; 305:251-254.
40. Wu TT, Johnson G, Kabat EA. Length distribution of CDRH3 in antibodies. Proteins 1993; 16:1-7.
41. Fleurant M, Changchien L, Chen CT et al. Shark Ig light chain junctions are as diverse as in heavy chains. J Immunol 2004; 173:5574-5582.
42. Wedemayer GJ, Patten PA, Wang LH et al. Structural insights into the evolution of an antibody combining site. Science 1997; 276:1665-1669.
43. Wilson IA, Stanfield RL. Antibody-antigen interactions: new structures and new conformational changes. Curr Opin Struct Biol 1994; 4:857-867.
44. Stanfield RL, Dooley H, Verdino P et al. Maturation of shark single-domain (IgNAR) antibodies: evidence for induced-fit binding. J Mol Biol 2007; 367:358-372.
45. Nguyen VK, Hamers R, Wyns L et al. Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire. EMBO J 2000; 19:921-930.
46. Achour I, Cavelier P, Tichit M et al. Tetrameric and homodimeric camelid IgGs originate from the same IgH locus. J Immunol 2008; 181:2001-2009.
47. Burnet FM. The clonal selection theory of acquired immunity. Cambridge: Cambridge University Press, 1959.
48. Stanhope-Baker P, Hudson KM, Shaffer AL et al. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. Cell 1996; 85:887-897.
49. Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. Annu Rev Immunol 1986; 4:339-368.
50. Alt FW, Yancopoulos GD, Blackwell TK et al. Ordered rearrangement of immunoglobulin heavy chain variable region segments. EMBO J 1984; 3:1209-1219.

51. Chowdhury D, Sen R. Regulation of immunoglobulin heavy-chain gene rearrangements. Immunol Rev 2004; 200:182-196.
52. Fuxa M, Skok J, Souabni A et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. Genes Dev 2004; 18:411-422.
53. Kosak ST, Skok JA, Medina KL et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science 2002; 296:158-162.
54. Jhunjhunwala S, van Zelm MC, Peak MM et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. Cell 2008; 133:265-279.
55. Roldán E, Fuxa M, Chong W et al. Locus "decontraction" and centromic recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. Nat Immunol 2005; 6:31-41.
56. Schlimgen RJ, Reddy KL, Singh H et al. Initiation of allelic exclusion by stochastic interaction of Tcrb alleles with repressive nuclear compartments. Nat Immunol 2008; 9:802-809.
57. Perlot T, Alt FW, Bassing CH et al. Elucidation of IgH intronic enhancer functions via germ-line deletion. Proc Natl Acad Sci USA 2005; 102:14362-4367.
58. Tunyaplin C, Knight KL. IgH gene rearrangements on the unexpressed allele in rabbit B-cells. J Immunol 1997; 158:4805-4811.
59. Lanning D, Jasper P, Knight K. IgH haplotype exclusion in rabbits. Semin Immunol 2002; 14:163-168.
60. Tunyaplin C, Knight KL. Fetal VDJ gene repertoire in rabbit: evidence for preferential rearrangement of VH1. Eur J Immunol 1995; 25:2583-2587.
61. Mage RG, Lanning D, Knight KL. B-cell and antibody repertoire development in rabbits: the requirement of gut-associated lymphoid tissues. Dev Comp Immunol 2006; 30:137-153.
62. Yancopoulos GD, Desiderio SV, Paskind M et al. Preferential utilization of the most JH-proximal VH gene segments in preB-cell lines. Nature 1984; 311:727-733.
63. Knight KL, Becker RS. Molecular basis of the allelic inheritance of rabbit immunoglobulin VH allotypes: implications for the generation of antibody diversity. Cell 1990; 60:963-970.
64. Reynaud CA, Imhof BA, Anquez V et al. Emergence of committed B-lymphoid progenitors in the developing chicken embryo. EMBO J 1992; 11:4349-4358.
65. Weill JC, Reynaud CA. The chicken B-cell compartment. Science 1987; 238:1094-1098.
66. Benatar T, Tkalec L, Ratcliffe MJ. Stochastic rearrangement of immunoglobulin variable-region genes in chicken B-cell development. Proc Natl Acad Sci USA 1992; 89:7615-7619.
67. Weill JC, Cocea L, Reynaud CA. Allelic exclusion: lesson from GALT species. Semin Immunol 2002; 14:213-215.
68. Du Pasquier L, Hsu E. Immunoglobulin expression in diploid and polyploid interspecies hybrid of Xenopus: evidence for allelic exclusion. Eur J Immunol 1983; 13:585-590.
69. Barreto V, Meo T, Cumano A. Mice triallelic for the Ig heavy chain locus: implications for VHDJH recombination. J Immunol 2001; 166:5638-5645.
70. Du Pasquier L, Blomberg B. The expression of antibody diversity in natural and laboratory-made polyploid individuals of the clawed toad Xenopus. Immunogenetics 1982; 15:251-60.
71. Hoegg S, Brinkmann H, Taylor JS et al. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol 2004; 59:190-203.
72. Lee V, Huang JL, Lui MF et al. The evolution of multiple isotypic IgM heavy chain genes in the shark. J Immunol 2008; 180:7461-7470.
73. Anderson M, Amemiya C, Luer C et al. Complete genomic sequence and patterns of transcription of a member of an unusual family of closely related, chromosomally dispersed Ig gene clusters in Raja. Int Immunol 1994; 6:1661-1670.
74. Malecek K, Lee V, Feng W et al. Immunoglobulin heavy chain exclusion in the shark. PLoS Biol 2008; 6:e157.
75. Eason DD, Litman RT, Luer CA et al. Expression of individual immunoglobulin genes occurs in an unusual system consisting of multiple independent loci. Eur J Immunol 2004; 34:2551-2558.
76. Mostoslavsky R, Singh N, Kirillov A et al. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. Genes Dev 1998; 12:1801-1811.
77. Mostoslavsky R, Singh N, Tenzen T et al. Asynchronous replication and allelic exclusion in the immune system. Nature 2001; 414:221-225.
78. Kokubo F, Litman R, Shamblott MJ et al. Diverse organization of immunoglobulin VH gene loci in a primitive vertebrate. EMBO J 1988; 7:3413-3422.
79. Ghaffari SH, Lobb CJ. Structure and genomic organization of a second cluster of immunoglobulin heavy chain gene segments in the channel catfish. J Immunol 1999; 162:1519-1529.
80. Reynaud CA, Dahan A, Anquez V et al. Somatic hyperconversion diversifies the single VH gene of the chicken with a high incidence on the D region. Cell 1989; 40: 283-291.

81. Lee SS, Fitch D, Flajnik MF et al. Rearrangement of immunoglobulin genes in shark germ cells. J Exp Med 2000; 191:1637-1648.
82. de Villartay J-P. Passera ou ne passera pas—accessibility is key. Nature Immunol 2006; 7:1019-1021.
83. Criscitiello MF, Flajnik MF. Four primordial immunoglobulin light chain isotypes, including lambda and kappa, identified in the most primitive living jawed vertebrates. Eur J Immunol 2007; 37:2683-2694.
84. Hsu E, Criscitiello MF. Diverse immunoglobulin light chain organizations in fish retain potential to revise B-cell receptor specificities. J Immunol 2006; 177:2452-2462.
85. Zimmerman AM, Yeo G, Howe K et al. Immunoglobulin light chain (IgL) genes in zebrafish: Genomic configurations and inversional rearrangements between (VL-JL-CL) gene clusters. Dev Comp Immunol 2008; 32: 421-434.
86. Bengtén E, Stromberg S, Daggfeldt A et al. Transcriptional enhancers of immunoglobulin light chain genes in Atlantic cod (Gadus morhua). Immunogenetics 2000; 51:647-658.
87. Nemazee D. Receptor editing in lymphocyte development and central tolerance. Nat Rev Immunol 2006; 6:728-740.
88. Danilova N, Bussmann J, Jekosch K et al. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. Nat Immunol 2005; 6:295-302.
89. Daggfeldt A, Bengtén E, Pilstrom L. A cluster type organization of the loci of the immunoglobulin light chain in Atlantic cod (Gadus morhua L.) and rainbow trout (Oncorhynchus mykiss Walbaum) indicated by nucleotide sequences of cDNAs and hybridization analysis. Immunogenetics 1993; 38:199-209.

Chapter 15

Normal and Pathological V(D)J Recombination:

Contribution to the Understanding of Human Lymphoid Malignancies

Saïda Dadi, Sandrine Le Noir, Vahid Asnafi, Kheïra Beldjord and Elizabeth A. Macintyre*

Abstract

The majority of haematological cancers involve the lymphoid system. They include acute lymphoblastic leukemias (ALL), which are arrested at variable stages of development and present with blood and bone marrow involvement and chronic leukemias, lymphomas and myelomas, which present with infiltration of a large variety of hematopoietic and non hematopoietic tissues by mature lymphoid cells which express a surface antigen receptor. The majority involve the B-cell lineage and the vast majority have undergone clonal rearrangement of their Ig and/or TCR rearrangements. Analysis of Ig/TCR genomic V(D)J repertoires by PCR based lymphoid clonality analysis within a diagnostic setting allows distinction of clonal from reactive lymphoproliferative disorders, clonal tracking for evidence of tumor dissemination and follow-up, identification of a lymphoid origin in undiagnosed tumors and evaluation of clonal evolution. Ig/TCR VDJ errors are also at the origin of recombinase mediated deregulated expression of a variety of proto-oncogenes in ALL, whereas in lymphoma it is increasingly clear that IgH containing translocations result from abnormalities other than VDJ errors (somatic hypermutation and/or isotype switching). In addition to this mechanistic contribution to lymphoid oncogenesis, it is possible that failure to successfully complete expression of an appropriate Ig or TCR may lead to maturation arrest in a lymphoid precursor, which may in itself contribute to altered tissue homeostasis, particularly if the arrest occurs at a stage of cellular expansion.

Introduction

Approximately 5% of human cancers overall and over 70% of haematological cancers involve the lymphoid system, with the majority involving the B-cell lineage. Lymphoid cancers include immature, "blastic" lymphoid proliferations which involve essentially the blood and/or bone marrow (Acute Lymphoblastic Leukemia or ALL), mature lympho-proliferations involving predominantly secondary lymphoid organs (non-Hodgkin's Lymphomas or NHL) or blood and bone marrow (chronic lymphocytic leukemias or CLL) and expansions of plasmocytes, with predominant bone marrow and tissue involvement (multiple myeloma or MM). Dysimmune states such as Hodgkin's disease or Angioimmunoblastic lymphadenopathy (AILD), at the interface

*Corresponding Author: Elizabeth A. MacIntyre—Hopital Necker Enfants Malades, Laboratoire d'hématologie, bât. Pasteur, 149 rue de Sèvres 75015 Paris, France.
Email: elizabeth.macintyre@nck.aphp.fr

V(D)J Recombination, edited by Pierre Ferrier. ©2009 Landes Bioscience and Springer Science+Business Media.

between reactive immune disorders and lymphoid malignancies, also exist. Whether these are classified as lymphoid malignancies or not often depend on the techniques available for their characterization, notably analysis of the V(D)J status of their immunoglobulin (Ig) and T-cell Receptor (TCR) loci by techniques which will be collectively referred to as lymphoid clonality analysis here. Lymphoid malignancies are also frequently characterized by V(D)J recombinase errors which lead to transcriptional deregulation of lymphoid "oncogenes" by juxtapositioning to, most commonly, Ig or TCR regulatory sequences. This represents a lymphoid specific form of "physiological genetic instability" which includes V(D)J recombinase errors and abnormalities of isotype switching and/or somatic mutation. Only the former will be considered here; they are collectively, if imprecisely, referred to as V(D)J translocations. Such errors can be considered to be, at a minimum, mechanistic elements involved in lymphoid oncogenesis. It is, however, possible that failure to successfully complete fabrication and expression of an appropriate Ig or TCR may in itself represent an oncogenic event within the multistage process that is now recognized to preceed clinical presentation of the majority of human cancers.

Since lymphoid cancers represent homogeneous populations arrested at different stages of development, they provide invaluable models for the study of molecular and cellular events leading to interruption of lymphoid development. Within this context, "reading the language" of Ig/TCR rearrangements can provide useful information regarding the type of lymphoid (sub)population involved, the stage of maturation arrest and the chromatin accessibility of the different Ig/TCR loci. It should however be emphasised that the pheno/genotype of the bulk lymphoid cancer is not necessarily synonymous, but most probably downstream, to the lymphoid cancer stem cell. Since any detectable clonal V(D)J rearrangement or translocation suggests at a minimum that the Ig/TCR loci were accessible during preceeding stages of lymphoid oncogenesis, such rearrangements represent useful fingerprints of upstream oncogenic events. We have undertaken to review these different, but interlinked, applications of the analysis of normal and abnormal V(D)J coding joint repertoires applied to understanding of lymphoid malignancies and their dysimmune close relatives. Such an approach is by definition nonexhaustive and we apologies to all individual contributors which we have only referenced indirectly, in the interests of brevity. We will not discuss therapeutic aspects of V(D)J manipulation, nor analysis of transcribed, functional V(D)J repertoires and will only briefly touch on detection of signal junction rearrangements.

Diagnostic Clonality Analysis

Molecular analysis of Ig/TCR genomic repertoires in diagnostic evaluation of (suspected) human lymphoid malignancies was initially performed by Southern blot analysis,[1-3] but was progressively replaced from the 1980s onwards by PCR analysis from DNA.[4-9] Both are based on the principal that reactive lymphoproliferations are associated with polyclonal Ig/TCR repertoires whereas the majority of lymphoid cancers demonstrate clonal, homogeneous rearrangements of Ig and/or TCR loci,[10] with the pattern of clonal rearrangements reflecting the lymphoid lineage involved and its stage of maturation arrest.[10-13]

Technical and Practical Aspects

Southern blotting predominantly reflected homogeneous V and J segment usage whereas PCR V(D)J amplification also exploits heterogeneity of VDJ junctional sequences at the third complementarity determining region (CDR3). The longer the CDR3, the easier the distinction of clonal and polyclonal rearrangements.[14] Detection of VDJ, DJ, VD DD and DJ rearrangements are possible if appropriate primers are used.[15] The majority of diagnostic systems use consensus primers directed to relatively conserved framework regions, often in a multiplex format.[15] Predictably, the risk of false negative results is dependent on the complexity of the repertoire (Table 1) and the degree of homology between the V, D and J primers and their target sequences. The other main factor contributing to false negativity is somatic mutation involving PCR primer target sequences but others include: presence of inhibitors; analysis of uninvolved tissue and DNA degradation of fixed tissues.

Table 1. Human Ig/TCR repertoires, combinatorial complexity and chromosomal localisation

	Number of Germline Encoded Segments					
	V	D	J	Approximate CDR3 Length(bp)	Number of N Regions	Chromosomal Localization
IgH	46-52	27	6	50	1-2	14q32.3
IgK	31-36	0	5	10	1	2p11.2
Igλ	30-33	0	4	10	1	22q11.2
TCRδ	7	3	4	5-50	1-4	14q11.2
TCRα	45-47	0	50	10	1	14q11.2
TCRγ	9	0	5	10	1	7q14
TCRβ	39-47	2	13	10	1-2	7q34

The number of V segments varies. Certain Vα/δ segments can rearrange to both TCRδ and TCRα loci. Number of N region varies with incomplete VD, DD or DJ rearrangements.

Distinction of clonal, oligoclonal and polyclonal PCR products is based on either nondenaturing polyacrylamide gel electrophoresis (PAGE), usually under conditions encouraging heteroduplex formation, or "genescan" sizing of fluorescent PCR products. The former has the advantage of optimising distinction of clonal homoduplexes from polyclonal heteroduplexes but requires optimal PAGE conditions. Genescan sizing allows precise information regarding clonal product size, useful for molecular follow-up and comparison of different samples from a given tumor and can allow identification of V and J segment usage if differently labelled primers are used (Fig. 1). Under qualitative conditions, both have an approximate sensitivity of 1-5%, although this depends on the position of clonal and polyclonal populations, since a clonal population which is situated at the peak of the Gaussian distribution of polyclonal PCR products will be detected with lower sensitivity than one which is either larger or smaller than these fragments (Fig. 1). Quantitation of clonal rearrangements by real-time PCR is possible using CRD3 specific probes, or more usually primers (Fig. 2) (ref. 16 and references there in) This requires sequencing of diagnostic material and has been developed essentially for follow-up of patients with ALL. In general, diagnostic strategies aim only to distinguish clonal from polyclonal populations and do not attempt to identify segment usage. Judicious use of appropriately situated, variably labelled fluorescent primers allows identification of V, D and J segments from a limited number of multiplex PCR, based on PCR product size and fluorescence.[17] "Reading the language" of Ig/TCR rearrangements in this way can contribute to identification of the stage of maturation arrest and lineage affiliation. Such analyses do not allow determination of functional, in-frame rearrangement, unless combined with sequence analysis.

Diagnostic PCR have been developed for all loci other than TCRα. The most widely used loci for diagnostic clonality analysis are IgH VDJ and TCRγ VJ, since both rearrange relatively early during normal B and T-lymphoid development respectively, including in all subsets of each lineage. Backup loci for the B-cell lineage include Igκ and IgH DJ rearrangements, whereas Igλ clonality analysis within a diagnostic setting is complex and rarely adds additional information. For suspected T-cell malignancies, TCRγ can be complemented by TCRβ VDJ analysis, which is a more appropriate target than TCRδ; due to the deletion of this locus during TCRα rearrangement and the consequent risk of pseudo-clonality from rare residual TCRδ rearrangements. Use of TCRδ is essentially restricted to clonality analysis in ALL and rare suspected TCRγδ lymphoproliferative disorders. Details regarding the incidence and patterns of Ig/TCR rearrangements in the main categories of lymphoproliferative disorders (LPD) can be found in Table 2.[15,18-24] Succinctly, mature B lineage LPD rearrange IgH and

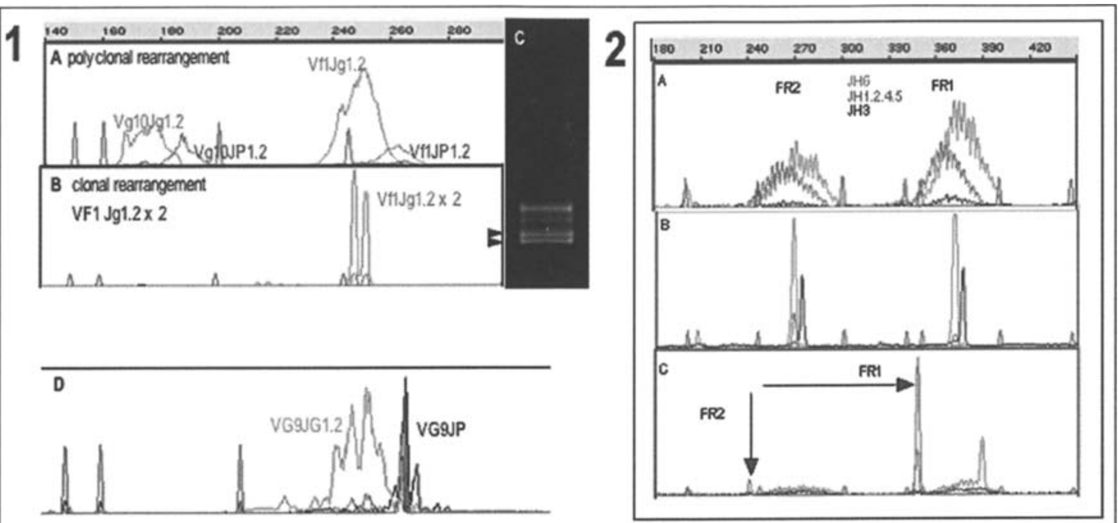


Figure 1. Lymphoid clonality analysis. Genescan and heteroduplex images of TCR and IgH genomic repertoires. 1)- Analysis of multiplex TCRγ rearrangement by PCR from DNA using fluorescent primers and genescan (GS) evaluation (left) compared to nondenaturing PAGE analysis of heteroduplex (HD) PCR products (right): A = polyclonal rearrangements. B and C - Clonal bi-allelic Vfl-Jγ1/2 rearrangements in a B-cell precursor ALL, analysed by GS (B) or HD (C). Slow migrating clonal heteroduplexes are seen when both alleles undergo rearrangement with the same Vγ and Jγ segments, but different CDR3. Homoduplexes are indicated by arrows. The use of differently labelled primers allows identification of the V and J segments used on the basis of PCR product size and colour, as shown. Size markers are in red. D = polyclonal TCR Vγ9-Jg rearrangement (green) and canonical Vγ9JP (black). The latter represents a selected, functional repertoire, as evidenced by the 3bp spacing between peaks, as seen with IgH, but not with other TCRγ rearrangements. 2)- Analysis of multiplex multifluorescent IgH rearrangement by PCR from DNA using fluorescent primers and genescan evaluation. Polyclonal (A) and clonal (B and C) rearrangements are shown. The use of 3 differently labelled JH primers (JH1,2,4,5 in blue, JH6 in green and JH3 in black) gives improved resolution in polyclonal rearrangements, the majority of which use JH4-6 segments. IgH rearrangements are amplified with both FR1 and FR2 consensus primers in separate reactions but analysed together, in order to increase clonal informativity in cases having undegone somatic mutation. For example, of the 2 B-cell precursor ALLs shown , sample B demonstrates a major JH6 and a minor JH3 rearrangement which are informative with both FR1 (right) and FR2 (left) VH primers. For sample C, in contrast, the major JH1,2,4,5 rearrangement is amplified efficiently from the FR1 but poorly from the FR2 primer and the minor JH6 rearrangement is only seen with FR1. A color version of this image is available at www.landesbioscience.com/curie. The uniform intensity, red peaks coorespond to size markers.

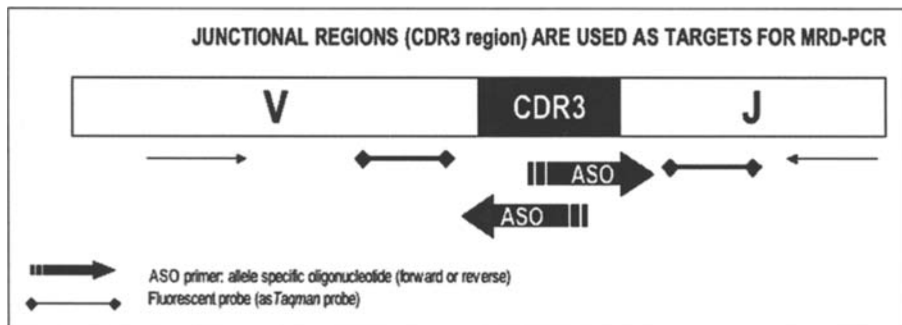


Figure 2. Real time quantitative Ig/TCR CDR3 specific strategies. Quantification by RQ-PCR of the tumor load or minimal residual disease (MRD). For each Ig/TCR rearrangement, the junctional region is amplified, sequenced and several "clone-specific" primers or allele specific oligonucleotides (ASO) are designed. Specific CDR3 specific primers are then used for clone specific amplification of follow-up material using CDR3 and V, D or J primers and V or J Taqman probes. Quantification is performed using a standard curve constructed from the RQ-PCR assay by serial dilutions of patient's blasts in a peripheral blood mononuclear cell pool (10^{-1} to 10^{-5}).

Igκ (VJ or Kappa deleting element—KDE) in the vast majority of cases, with extensive repertoires which lead to little risk of false positive results and a risk of false negative results which is proportional to the degree of somatic mutation. Mature T lineage LPD rearrange TCRγ and TCRβ and occasionally TCRδ. The restricted repertoire of TCRγ VJ rearrangements leads to a risk of false positive detection of pseudoclonality, particularly if PAGE conditions are suboptimal.[15] PAGE heteroduplex analysis is preferable to fluorescent genescan analysis in a diagnostic setting, since there is a lower risk of false positives. The presence of canonical "invariant" rearrangements, such as Vγ9-JP rearrangements in circulating TCRγδ lymphocytes, can also be erroneously interpreted as indicating clonal expansion by inexperienced operators and for this reason not all diagnostic multiplex strategies include a JP (also referred to as Jγ1.2) specific primer.[15] Two classifications for human TCRγ V and J segments exist.[15] The presence of minor normal clonal/invariant populations is well recognised in circulating CD8+ T-lymphocytes from older individuals and in reactive disorders such as lymphomatoid papulosis. The risk of false positive results can be minimised by simultaneous use of TCRβ analysis[19] and restriction of these analyses to high throughput laboratories, in order to maximise experience. Interpretation of lymphoid clonality profiles should be undertaken in close interaction with the prescribing physician or pathologist and with knowledge of the clinical context.

Cross lineage rearrangements, also referred to as "illegitimate rearrangements" (Ig rearrangements in a T LPD or vice versa) are rare in mature LPD. They are common in acute lymphoblastic leukemias, with the majority of B lineage ALL demonstrating TCRγ rearrangement and/or TCRδ or, more rarely TCRβ rearrangements. Ig rearrangements in T-ALL are less common and are preferentially found in the TCRγδ lineage.[25-27] Illegitimate rearrangements reflect the fact that ALLs remain recombinase competent and consequently rearrange all loci in an accessible chromatin configuration. Rearrangement patterns differ with oncogenic subtype, with stage of maturation arrest and with patient age. Amongst B-cell precursor ALLs, for example, relatively mature cases which express Ig cyt μ rarely demonstrate TCRγ rearrangements, whereas the majority of $CD10^+$ cyt μ- ETV6-RUNX1 or BCR-ABL cases do so.[28,29] Details of Ig/TCR rearrangement profiles in ALL can be found in Table 2. Detection of lymphoid clonality is rarely required to make a diagnosis of ALL and is essentially used for molecular follow-up (see below).[16] Extensive sequence analysis of these clonal rearrangements has, however, allowed accumulation of a large databank allowing analysis of V, D and J segment usage and CDR3 diversity, which may eventually lead to improved understanding of the pathogenic stages leading to ALL development.

Table 2. Approximate incidence of clonal Ig/TCR rearrangement in lymphoid malignancies, as detected by PCR from DNA. Only diagnostic PCR targets are cited

	IgH	IgK	Igλ	TCRδ	TCRγ	TCRβ
B-Cell Proliferation						
BCP-ALL	90	30	20	30	60	30
CLL	100	100	30	10	20	25
Non-Hodgkin's Lymphoma						
FL	90	85	20	5	5	5
MCL	100	100	45	5	10	10
BLBCL	85	80	30	15	15	20
MZL	95	80	30	10	15	20
T-Cell Proliferation						
T-ALL	5	0	0	50	90	90
T-LGL	0	5	5	30	95	95
AILT	30	30	5	35	90	90

Abbreviations: BCP-ALL: B-cell precursor Acute Lymphoblastic Leukemia; CLL: Chronic Lymphocytic leukemia; FL: Follicular Lymphoma; MCL: Mantle Cell Lymphoma; BLBCL: Diffuse Large B-Cell Lymphoma; MZL: Marginal Zone Lymphoma; T-ALL: T-cell acute lymphoblastic leukaemia; T-LGL: T-Large Granular lymphocytic leukaemia; AILT: Angioimmunoblastic T-Cell Lymphoma.

Clinical Applications

Diagnostic clonality analysis is mainly used to distinguish reactive, polyclonal LPD from clonal, probably but not necessarily, malignant LPD. Once a clonal population has been identified, it is possible to track this clone in different tissue samples, in order to assess dissemination at diagnosis, or to determine clonal identity at relapse. Clonal tracking has also been used within a minimal residual disease setting in ALL and certain NHL, once apparent complete remission has been obtained, to stratify individual patient management, based on the cinetics of response to remission induction at diagnosis. Succinctly, clonal tracking with CDR3 specific probes, used within a strictly standardised, quantitative setting, allow the detection of minor clonal poulations with a reproducible sensitivity of at least 10^{-4} (1 malignant cell amongst 10000 normal cells). It has also been used to "back-track" preclinical development of ALL, in conjuction with molecular oncogenic markers, allowing the identification of leukemic clones many years before clinical presentation, including in postnatal samples prior to development of pediatric ALL.[30,31]

Recombinase Mediated Oncogenesis

Analysis of structural chromosomal abnormalities by classical morphological karyotyping in lymphoid malignancies allowed the identification of recurrent translocations involving the Ig loci in B lymphoid malignancies and TCR loci in T-cell malignancies. The advent of molecular techniques led to identification of the Ig/TCR partner genes and the demonstration that karyotypic analysis largely underestimated the incidence and complexity of these rearrangements. The large number of partner genes identified has allowed numerous insights into normal and pathological lymphoid development and function, but their very number precludes their description here and readers are invited to consult the following reviews on the subject.[32-36] Only general aspects relevant to V(D)J rearrangement will be detailed here. Within the context of lymphoid

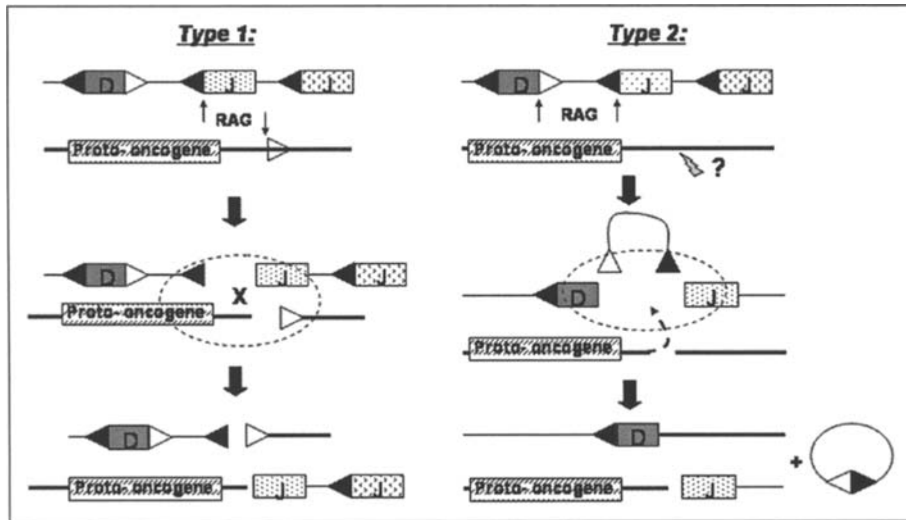


Figure 3. Type 1 and 2 Ig/TCR rearrangements. Recognition signal sequences (RSS) are represented by triangles at V, D and J segment extremities. Gray triangles represent cryptic RSS. In Type I rearrangements, RAG targets both the bona-fide and the cryptic RSS. In Type 2 junctions, the break in the proto-oncogene is targeted by unknown mechanisms. In both cases, there is excision of intervening DNA, in the signal joint shown at the bottom right hand corner.

malignancies, the term "illegitimate" rearrangement is usually reserved for cross-lineage intralocus rearrangements, such as the TCR rearrangements identified in B lineage ALL described above. "Trans-rearrangement" refers to rearrangement between distinct Ig and TCR loci, abnormalities which have been principally described in patients with Ataxia Telangiectasia.[37] V(D)J translocations usually implies structural karyotypic abnormalities involving a proto-oncogene and an Ig or TCR locus (Fig. 3 and Table 3). The increasing recognition of recombinase mediated deregulation of genes with no involvement of an Ig/TCR locus, including those resulting from microscopic, intragenic rearrangements not associated with evident karytoypic abnormalities, justifies use of the more general term "recombinase mediated oncogenesis". Comparative genomic hybridization has demonstrated that in pediatric B lineage ALL, many of these deletions involve genes which regulate B-cell development, including TCF3 (also known as E2A), EBF1, LEF1, IKZF1 (IKAROS) and IKZF3 (AIOLOS).[38] At least a proportion of these are mediated by the recombinase.

Recombinase mediated events can occur at the site of any RSS-like sequence which is in an accessible chromatin configuration during recombinase activity. One of the best studied examples outside the lymphoid oncogenesis context is deletions of the HPRT locus.[39-42] These have been used as a measure of genomic instability, some of which are mediated by the recombinase complex. Within the present context, only V(D)J recombinase mediated events with oncogenic potential will be detailed. The role of recombinase abnormalities in Ig/TCR rearrangements is illustrated by their high incidence in patients with Ataxia Telangiectasia and similar disorders.[43]

VDJ Errors in Lymphoid Malignancies

During lymphoid development, recombinase activity targeted to recombination signal sequences (RSS) would ideally be restricted to legitimate targets within Ig/TCR loci and all genes controlling tissue homeostasis would be protected from this lymphoid specific form of "physiological genomic instability". The existence, however, of a large number of RSS-like sequences throughout the genome (10 million or 1 cryptic RSS every 1-2 kb on average) means that nonspecific targeting of RAG1 can induce double stranded breaks outside Ig/TCR loci, leading to

Table 3. Deregulation of lymphoid oncogenes by Ig/TCR juxtapositioning in ALL

Oncogene (Ig/tcr Partner Genes)	Protein Family Group	Translocations Involved	References
B-ALL	**Translocation Involving Ig Genes**		
ID4	Inhibitor of DNA binding(ID)HLH	t(6; 14)(p21; q32)	61
LHX4	LIM-homeodomain	t(1; 14)(q25; q32)	60
BCL9	Not identified	t(1; 14)(q21; q32)	59
IL3	4H Cytokine	t(5; 14)(q32; q32)	62
c-Myc	bHLH-Zip	t(8; 14)(q24; q32)	58,55
		t(2; 8)(p12; q24)	56
		t(8; 22)(q24; q11)	57
CEBP	bZIP	t(14; 19)(q32; q13);	63
		t(8; 14)(q11; q32);	
		Inv(14)(q11; q32)/	
		t(14; 14)(q11; q32)	
		t(14; 20)(q32; q13)	
T-ALL	**Translocation Involving TCR Genes**		
HOXA cluster	Class I homeodomaincontaining	Inv(7)(p15q34)t(7; 7)	90,91
TLXI(HOXII)	Class II homeodomaincontaining	t(7; 10)(q34; q24)	65,66
		t(10; 14)(q24; q11)	
TLX3 (HOXIIL2)*	Class II homeodomaincontaining	t(5; 14)(q35; q32)	86,87
LMO1	LIM-only domain	t(11; 14)(p15; q11)	68
LMO2	LIM-only domain	t(11; 14)p13; q11),	69
		t(7; 11)(q35; q13)	71
TAL1	b HLH Type II	t(1; 14)(p32; q11),	100
		t(1; 7)(p32; q34)	91
TAL2	b HLH Type II	t(7; 9)(q34; q32)	75
LCK	SRC family of tyrosine kinase	t(1; 7)(p34; q34)	95,101
BHLHB1	b HLH Type II	t(14; 21)(q11.2; q22)	76
LYL1	b HLH Type II	t(7; 19)(q34; p13)	74
CCND2	D-type cyclin	t(7; 12)(q34; p13)	102
		t(12; 14)(p13; q11)	103
NOTCH1	Notch receptor family	t(7; 9)(q34; q34.3)	94

**TLX3* is included despite the fact that the predominant t(5; 14) involves BCL11B, not IgH, since these *BCL11B-TLX3* translocations are mediated by the recombinase and since rare translocations involving *TLX3* and *TCRα/δ* are described.

intergenic rearrangements and deregulation of genes by juxtapositioning to Ig/TCR regulatory sequences (promoters or enhancers).[44] This can lead to increased expression or nonextinction of the juxtaposed "proto-oncogene" by promoter/enhancer substitution or by separation of coding sequences from negative regulatory elements. Only those rearrangements which lead to deregulated tissue homeostasis will be associated with lymphoid malignancies. If the deregulated genes induce a survival or proliferative advantage or a block to maturation, the clone bearing the translocation will be transformed, or at least immortalised. Based on these considerations, V(D)J errors will only occur in cells which are recombinase competent and will target proto-oncogenes which are accessible during this phase of recombinase activity.

It is increasingly recognised that the transcriptional and phenotypic profile observed in a cancer at diagnosis is not necessarily identical, but is probably more mature, when compared to the cancer initiating or stem cell. Genetic modifications which occur in this cancer stem cell are, however, transmitted to all clonal descendants. Within this context, both bona-fide Ig/TCR rearrangements and recombinase mediated oncogenic rearrangements detected in diagnostic material can represent genetic fingerprints of earlier events which have occurred in lymphoid cancer stem cells, or in intermediate malignant precursor populations. If such markers are present in the majority of the tumor at diagnosis, it is likely that they reflect an upstream event during oncogenic development, wheras those present in minor subclones are more likely to represent downstream events occurring in tumor subclones. The capacity to accurately evaluate the proportion of cells demonstrating a given marker depends on the techniques used. Briefly, molecular PCR and CGH based techniques using extracted DNA are poorly adapted to precise quantification and cytogenetic analysis of mitotic material is biased by potential nonrepresentativity of the cells undergoing mitosis under the culture conditions used. FISH analysis of interface nuclei has the advantage of being cell based, but is only applicable to certain oncogenic markers, not to V(D)J rearrangements and is heavily dependent on the quality of material analysed (bare nuclei vs. tissue sections, for example). Given these reserves, detection of an Ig or TCR rearrangement in an apparently nonlymphoid cancer, implies prior exposure of malignant precursors to recombinase activity. Identification of Ig/TCR rearrangements in Acute Myeloid leukaemia, for example, is preferentially found in cases with MLL gene rearrangement, with the MLL fusion transcript partners being associated with different Ig/TCR profiles.[45] Similarly, detection of a recombinase mediated oncogenic marker implies chromatin accessibility of the partner gene during a phase of recombinase competence prior to tumor development. What level of qualitative and/or quantitative recombinase competence and/or RAG1/2 activity is required for these recombinase errors is not clear. Rearrangement of TCRδ and TCRγ can occur in the presence of much lower levels of RAG1 activity than that required for TCRβ rearrangement [46] and it is possible to induce TCRδ rearrangement in kidney cells in the presence of E2A and HEB.[47,48]

Categories of Recombinase Errors

Two categories of recombinase errors are recognised:[44,49,50] Type I rearrangements demonstrate breaks at RSS at both loci, one of which is usually an Ig or TCR; in Type II rearrangements, only the Ig/TCR break is mediated by RAG and the mechanisms targeting the double stranded break on the partner gene are incompletely understood (Fig. 3). Once generated, this DNA fragment becomes included in the recombinase complex, with the translocation resulting from a DNA repair error, rather than mistargetting of the recombinase. A recombinase mediated error is characterized by i) involvement of an Ig/TCR locus; ii) recurrent genomic breakpoints; iii) identification of a bona-fide RSS-like sequence at the breakpoint on the partner chromosome iv) addition of nongermline encoded nucleotides at the translocation breakpoint and v) generation of a signal joint. Recombinase mediated translocations were first identified in B lymphoid non-Hodgkin's lymphoma (NHL) with the t(14; 18) translocation involving IgH and BCL2.[51-53] Translocations involving Ig loci preferentially involve the IgH locus and are found in relatively mature, sIg+ lymphomas. These translocations are essentially Type II and primarily involve abnormalities of class switch and somatic hypermutation; [54] as such, they are beyond the scope of this article, which is restricted to V(D)J recombinase errors

in immature lymphoproliferative disorders, essentially ALL. A proportion of these abnormalities are also found in certain lymphomas, notably those involving *MYC* in Burkitt's lymphoma and those involving *HOX11/TLX1* in T-lymphoblastic lymphoma.

VDJ Deregulation with Oncogenic Potential

Ig translocations are found in approximately 1% of B lineage ALL, when they are virtually restricted to mature, sIg + cases. Partner genes include MYC,[55-58] BCL-9,[59] LHX4,[60] ID4,[61] IL3[62] or the different members of the CEBP family[63] (Table 3). In contrast, chromosomal abnormalities involving the TCR loci are among those most frequently encountered in T-ALL. Most involve the TCRα/δ locus on chromosome 14q11 or more rarely, TCRβ on chromosome 7q34; [35,64] rearrangements involving TCRγ are exceptional. The first TCR translocations to be described in T-ALL were those involving HOX11/TLX1 at chromosome 10q24[65-67] and LMO1/2 on chromosome 11p.[68-71] The incidence of TCR translocations by classical, morphological karyotyping was underestimated and it was only with the advent of screening by FISH that the true incidence was appreciated. Screening for TCR translocations demonstrated that approximately 45% of T-ALLs demonstrate translocations, including a minority with as yet unidentified partners. Predominant known TCR partner genes can be divided into those of the bHLH, LMO and HOX/TLX families.

The most commonly encountered bHLH partner is TAL1/SCL, which was initially described in the rare t(1; 14)(p32; q11).[72] Much more frequent is the SIL-TAL1 recombinase mediated intrachromosomal deletion, which places the entire TAL1 coding sequence under control of the SIL promoter.[73] SIL-TAL1 deletions are found in 20% of pediatric and 5-10% of adult T-ALLs. Other bHLH translocations include the rare t(7; 19)(q34; p13),[74] t(7; 9)(q34; q32)[75] and t(14; 21)(q11; q22)[76] involving LYL-1, TAL2 and bHLHB1 respectively. The frequent involvement of members of the bHLH family of transcriptional regulators is coherent with the fundamental role of bHLH proteins in regulation of T and B lymphoid lineage development . This is further emphasised by the fact that the LMO proteins deregulated by TCR juxtapositioning in translocations involving LMO1 (11p15)[68] or LMO2 (11p13)[69,71] form part of a complex which also includes TAL1 and its bHLH partner, E2A.[77-81]

Deregulation of homeobox gene expression is increasingly recognised in T-ALL. The orphan homeobox gene, HOX11/TLX1, is predominantly involved in the t(10; 14)(q24; q11) and more rarely the t(7; 10)(q34; q24).[65-67] Forced expression of TLX1 in murine bone marrow gives rise to T-ALL-like malignancies with long latency, suggesting that other events are necessary to induce leukemia; but with TLX1 expression representing an early event.[82,83] TLX1 regulates the G1/S checkpoint of T-ALL via its biding capability to the protein serine/threonine phosphatases PP2A and PP1.[84,85] Chromosomal translocations t(10; 14)(q24; q11) involving TLX1 are amongst the clearest example of recombinase involvement in T-ALL. Deregulated expression of HOX11L2/TLX3 is frequently found in pediatric T-ALL, due in most cases to a t(5; 14) involving the TLX3 locus at 5q35 and CTIP2/ BCL11B at 14q32, 7000 kb proximal to the IgH locus.[86,87] Despite the absence of Ig/TCR involvement, this translocation is mediated by the recombinase and rare translocations involving TLX3 and TCRα/δ have been described.[88] TLX3 has very close homology to TLX1, as evidenced by microarray studies showing that TLX1 and TLX3 T-ALLs cluster together.[89,90] TCR translocations involving the HOXA cluster on chromosome 7 predominantly involve the TCRβ locus, leading to a cryptic intrachromosomal inversion.[90,91] Another common abnormality in T-ALL is deletion of the p16/INK4/Cdk2 gene; [92,93] this is recombinase mediated in at least a proportion of cases. Other rare recombinase mediated abnormalities include translocations involving TCRβ and Notch1 in the t(7; 9)(q34; q34)[94] and t(1; 7)(q34; q34) involving LCK and TCRβ.[95,96]

In general, these recombinase mediated errors are restricted to T-ALLs of the TCRαβ lineage, which express RAG1 and have undergone extensive TCR rearrangement. The HOX/TLX cases are arrested prior to TCRα rearrangement, in contrast to SIL-TAL1 cases, which have undergone TCRα rearrangement on at least one allele. They are rarely found in TCRγδ expressing T-ALLs, with the exception of TLX3 expressing cases, which frequently express both

TCRγδ and cytoplasmic TCRβ.[97] Recombinase V(D)J errors are clearly an important mechanism in the development of immature T-cell malignancies. Attempts to recreate these malignancies in murine models have frequently demonstrated long latency and/or a low proportion of mice developing leukaemia/lymphoma, in keeping with multistage oncogenesis.[98,99] In keeping with this, low levels of translocations involving LMO2 have been identified in normal thymus.[49]

Conclusion

The aforementioned abnormalities represent a mechanistic role for the recombinase in development of lymphoid malignancies. It is however also possible that failure to complete production of a mature, functional Ig or TCR may favour malignant expansion, particularly if the cells are arrested at a stage when the pre B or TCR is expressed and capable of mediating ligand driven cellular expansion. The majority of acute leukaemias do not express a surface Ig/TCR, despite having undergone extensive Ig/TCR rearrangement. This failure to complete successful Ig/TCR rearrangement is likely to be at least partially at the origin of the recombinase competence and the maintenance of RAG1 expression. It is therefore at least theoretically possible that abrogation of the factors blocking completion of Ig or TCR assembly could lead to expression of the appropriate Ig/TCR at the surface, downregulation of RAG expression and possibly even leukemic cell death by differentiation. As mentioned above, a significant proportion of HOX11L2/TLX3+ T-ALLs express unusual TCRγδ receptors and cytoplasmic TCRβ. These T-ALLs maintain high levels of RAG1 transcripts, despite the expression of a surface TCR, suggesting that expression of an "inappropriate, default" TCRγδ in cells having undergone beta selection is insufficient to allow extinction of the recombinase. Exploration of the mechanisms underlying the failure to rearrange TCRα may further our understanding of T-ALL oncogenesis.

In conclusion, understanding and exploiting normal and abnormal recombinase activity can be used both in individual patient management and in understanding lymphoid oncogenesis.

Acknowledgements

We thank all the technicians and students from the immunogenetics section of the diagnostic haematology platform and INSERM EMIU 0210 at Necker-Enfants Malades who have contributed over the years to accumulation of our experience of lymphoid clonality analysis. Thanks also to Claudine Schiff, Bertrand Nadel and Pierre Ferrier at Centre d'Immunologie Marseille Luminy for fruitful discussions and collaborations and to Jacques van Dongen et al for the coordination of the Biomed2 programs which allowed standardisation and optimised exploitation of clonality analysis in lymphoproliferative disorders. We also thank all French adult and pediatric clinicians who manage patients with lymphoid malignancies, without whose collaboration it would be impossible to further our understanding of these disorders.

Note

Saïda Dadi and Sandrine Le Noir have contributed equally to this work.

References

1. Beishuizen A, Verhoeven MA, Mol EJ et al. Detection of immunoglobulin heavy-chain gene rearrangements by Southern blot analysis: recommendations for optimal results. Leukemia 1993; 7(12):2045-2053.
2. Langerak AW, Wolvers-Tettero IL, van Dongen JJ. Detection of T-cell receptor beta (TCRB) gene rearrangement patterns in T-cell malignancies by Southern blot analysis. Leukemia 1999; 13(6):965-974.
3. van Krieken JH, Elwood L, Andrade RE et al. Rearrangement of the T-cell receptor delta chain gene in T-cell lymphomas with a mature phenotype. Am J Pathol 1991; 139(1):161-168.
4. Langerak AW, Szczepanski T, van der Burg M et al. Heteroduplex PCR analysis of rearranged T-cell receptor genes for clonality assessment in suspect T-cell proliferations. Leukemia 1997; 11(12):2192-2199.
5. Gonzalez M, Gonzalez D, Lopez-Perez R et al. Heteroduplex analysis of VDJ amplified segments from rearranged IgH genes for clonality assessments in B-cell non-Hodgkin's lymphoma. A comparison between different strategies. Haematologica 1999; 84(9):779-784.
6. Derksen PW, Langerak AW, Kerkhof E et al. Comparison of different polymerase chain reaction-based approaches for clonality assessment of immunoglobulin heavy-chain gene rearrangements in B-cell neoplasia. Mod Pathol 1999; 12(8):794-805.

7. Krafft AE, Taubenberger JK, Sheng ZM et al. Enhanced sensitivity with a novel TCRgamma PCR assay for clonality studies in 569 formalin-fixed, paraffin-embedded (FFPE) cases. Mol Diagn 1999; 4(2):119-133.
8. Assaf C, Hummel M, Dippel E et al. High detection rate of T-cell receptor beta chain rearrangements in T-cell lymphoproliferations by family specific polymerase chain reaction in combination with the GeneScan technique and DNA sequencing. Blood 2000; 96(2):640-646.
9. Theriault C, Galoin S, Valmary S et al. PCR analysis of immunoglobulin heavy chain (IgH) and TcR-gamma chain gene rearrangements in the diagnosis of lymphoproliferative disorders: results of a study of 525 cases. Mod Pathol 2000; 13(12):1269-1279.
10. van Dongen JJ. Analysis of immunoglobulin genes and T-cell receptor genes as a diagnostic tool for the detection of lymphoid malignancies. Neth J Med 1987; 31(5-6):201-209.
11. Gleissner B, Maurer J, Thiel E. Detection of immunoglobulin heavy chain genes rearrangements in B-cell leukemias, lymphomas, multiple myelomas, monoclonal and polyclonal gammopathies. Leuk Lymphoma 2000; 39(1-2):151-155.
12. Sen F, Vega F, Medeiros LJ. Molecular genetic methods in the diagnosis of hematologic neoplasms. Semin Diagn Pathol 2002; 19(2):72-93.
13. Langerak AW, van Krieken JH, Wolvers-Tettero IL et al. The role of molecular analysis of immunoglobulin and T-cell receptor gene rearrangements in the diagnosis of lymphoproliferative disorders. J Clin Pathol 2001; 54(7):565-567.
14. Miqueu P, Guillet M, Degauque N et al. Statistical analysis of CDR3 length distributions for the assessment of T and B-cell repertoire biases. Mol Immunol 2007; 44(6):1057-1064.
15. van Dongen JJ, Langerak AW, Bruggemann M et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2003; 17(12):2257-2317.
16. van der Velden VH, Cazzaniga G, Schrauder A et al. Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. Leukemia 2007; 21(4):604-611.
17. Delabesse E, Burtin ML, Millien C et al. Rapid, multifluorescent TCRG Vgamma and Jgamma typing: application to T-cell acute lymphoblastic leukemia and to the detection of minor clonal populations. Leukemia 2000; 14(6):1143-1152.
18. van Krieken JH, Langerak AW, San Miguel JF et al. Clonality analysis for antigen receptor genes: preliminary results from the Biomed-2 concerted action PL 96-3936. Hum Pathol 2003; 34(4):359-361.
19. Droese J, Langerak AW, Groenen PJ et al. Validation of BIOMED-2 multiplex PCR tubes for detection of TCRB gene rearrangements in T-cell malignancies. Leukemia 2004; 18(9):1531-1538.
20. Evans PA, Pott C, Groenen PJ et al. Significantly improved PCR-based clonality testing in B-cell malignancies by use of multiple immunoglobulin gene targets. Report of the BIOMED-2 Concerted Action BHM4-CT98-3936. Leukemia 2007; 21(2):207-214.
21. van Krieken JH, Langerak AW, Macintyre EA et al. Improved reliability of lymphoma diagnostics via PCR-based clonality testing: report of the BIOMED-2 Concerted Action BHM4-CT98-3936. Leukemia 2007; 21(2):201-206.
22. Bruggemann M, White H, Gaulard P et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia 2007; 21(2):215-221.
23. Langerak AW, Molina TJ, Lavender FL et al. Polymerase chain reaction-based clonality testing in tissue samples with reactive lymphoproliferations: usefulness and pitfalls. A report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 2007; 21(2):222-229.
24. Liu H, Bench AJ, Bacon CM et al. A practical strategy for the routine use of BIOMED-2 PCR assays for detection of B- and T-cell clonality in diagnostic haematopathology. Br J Haematol 2007; 138(1):31-43.
25. Asnafi V, Beldjord K, Garand R et al. IgH DJ rearrangements within T-ALL correlate with cCD79a expression, an immature/TCRgammadelta phenotype and absence of IL7Ralpha/CD127 expression. Leukemia 2004; 18(12):1997-2001.
26. Kitchingman GR, Rovigatti U, Mauer AM et al. Rearrangement of immunoglobulin heavy chain genes in T-cell acute lymphoblastic leukemia. Blood 1985; 65(3):725-729.
27. Szczepanski T, Pongers-Willemse MJ, Langerak AW et al. Ig heavy chain gene rearrangements in T-cell acute lymphoblastic leukemia exhibit predominant DH6-19 and DH7-27 gene usage, can result in complete V-D-J rearrangements and are rare in T-cell receptor alpha beta lineage. Blood 1999; 93(12):4079-4085.
28. Chen Z, Le Paslier D, Dausset J et al. Human T-cell gamma genes are frequently rearranged in B-lineage acute lymphoblastic leukemias but not in chronic B-cell proliferations. J Exp Med 1987; 165(4):1000-1015.
29. Brumpt C, Delabesse E, Beldjord K et al. The incidence of clonal T-cell receptor rearrangements in B-cell precursor acute lymphoblastic leukemia varies with age and genotype. Blood 2000; 96(6):2254-2261.

30. Greaves M. Molecular genetics, natural history and the demise of childhood leukaemia. Eur J Cancer 1999; 35(14):1941-1953.
31. Greaves M. Childhood leukaemia. BMJ 2002; 324(7332):283-287.
32. Graux C, Cools J, Michaux L et al. Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast. Leukemia 2006; 20(9):1496-1510.
33. O'Neil J, Look AT. Mechanisms of transcription factor deregulation in lymphoid cell transformation. Oncogene 2007; 26(47):6838-6849.
34. Argiropoulos B, Humphries RK. Hox genes in hematopoiesis and leukemogenesis. Oncogene 2007; 26(47):6766-6776.
35. Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. J Clin Oncol 2005; 23(26):6306-6315.
36. Willis TG, Dyer MJ. The role of immunoglobulin translocations in the pathogenesis of B-cell malignancies. Blood 2000; 96(3):808-822.
37. Stern MH, Lipkowitz S, Aurias A et al. Inversion of chromosome 7 in ataxia telangiectasia is generated by a rearrangement between T-cell receptor beta and T-cell receptor gamma genes. Blood 1989; 74(6):2076-2080.
38. Mullighan CG, Goorha S, Radtke I et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 2007; 446(7137):758-764.
39. Fuscoe JC, Zimmerman LJ, Lippert MJ et al. V(D)J recombinase-like activity mediates hprt gene deletion in human fetal T-lymphocytes. Cancer Res 1991; 51(21):6001-6005.
40. Fuscoe JC, Zimmerman LJ, Harrington-Brock K et al. Deletion mutations in the hprt gene of T-lymphocytes as a biomarker for genomic rearrangements important in human cancers. Carcinogenesis 1994; 15(7):1463-1466.
41. Fuscoe JC, Zimmerman LJ, Harrington-Brock K et al. Large deletions are tolerated at the hprt locus of in vivo derived human T-lymphocytes. Mutat Res 1992; 283(4):255-262.
42. Finette BA, Kendall H, Vacek PM. Mutational spectral analysis at the HPRT locus in healthy children. Mutat Res 2002; 505(1-2):27-41.
43. Digweed M. Human genetic instability syndromes: single gene defects with increased risk of cancer. Toxicol Lett 1993; 67(1-3):259-281.
44. Marculescu R, Vanura K, Montpellier B et al. Recombinase, chromosomal translocations and lymphoid neoplasia: targeting mistakes and repair failures. DNA Repair (Amst) 2006; 5(9-10):1246-1258.
45. Dupret C, Asnafi V, Leboeuf D et al. IgH/TCR rearrangements are common in MLL translocated adult AML and suggest an early T/myeloid or B/myeloid maturation arrest, which correlates with the MLL partner. Leukemia 2005; 19(12):2337-2338.
46. Asnafi V, Beldjord K, Boulanger E et al. Analysis of TCR, pT alpha and RAG-1 in T-acute lymphoblastic leukemias improves understanding of early human T-lymphoid lineage commitment. Blood 2003; 101(7):2693-2703.
47. Langerak AW, Wolvers-Tettero IL, van Gastel-Mol EJ et al. Basic helix-loop-helix proteins E2A and HEB induce immature T-cell receptor rearrangements in nonlymphoid cells. Blood 2001; 98(8):2456-2465.
48. Ghosh JK, Romanow WJ, Murre C. Induction of a diverse T-cell receptor gamma/delta repertoire by the helix-loop-helix proteins E2A and HEB in nonlymphoid cells. J Exp Med 2001; 193(6):769-776.
49. Marculescu R, Le T, Simon P et al. V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites. J Exp Med 2002; 195(1):85-98.
50. Lieber MR, Yu K, Raghavan SC. Roles of nonhomologous DNA end joining, V(D)J recombination and class switch recombination in chromosomal translocations. DNA Repair (Amst) 2006; 5(9-10):1234-1245.
51. Tsujimoto Y, Gorham J, Cossman J et al. The t(14; 18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. Science 1985; 229(4720):1390-1393.
52. Cleary ML, Smith SD, Sklar J. Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14; 18) translocation. Cell 1986; 47(1):19-28.
53. Tsujimoto Y, Croce CM. Analysis of the structure, transcripts and protein products of bcl-2, the gene involved in human follicular lymphoma. Proc Natl Acad Sci USA 1986; 83(14):5214-5218.
54. Schlissel MS, Kaffer CR, Curry JD. Leukemia and lymphoma: a cost of doing business for adaptive immunity. Genes Dev 2006; 20(12):1539-1544.
55. Dalla-Favera R, Martinotti S, Gallo RC et al. Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. Science 1983; 219(4587):963-967.
56. Davis M, Malcolm S, Rabbitts TH. Chromosome translocation can occur on either side of the c-myc oncogene in Burkitt lymphoma cells. Nature 1984; 308(5956):286-288.
57. Hollis GF, Mitchell KF, Battey J et al. A variant translocation places the lambda immunoglobulin genes 3′ to the c-myc oncogene in Burkitt's lymphoma. Nature 1984; 307(5953):752-755.

58. Manolov G, Manolova Y. Marker band in one chromosome 14 from Burkitt lymphomas. Nature 1972; 237(5349):33-34.
59. Willis TG, Zalcberg IR, Coignet LJ et al. Molecular cloning of translocation t(1; 14)(q21; q32) defines a novel gene (BCL9) at chromosome 1q21. Blood 1998; 91(6):1873-1881.
60. Kawamata N, Sakajiri S, Sugimoto KJ et al. A novel chromosomal translocation t(1; 14)(q25; q32) in preB acute lymphoblastic leukemia involves the LIM homeodomain protein gene, Lhx4. Oncogene 2002; 21(32):4983-4991.
61. Bellido M, Aventin A, Lasa A et al. Id4 is deregulated by a t(6; 14)(p22; q32) chromosomal translocation in a B-cell lineage acute lymphoblastic leukemia. Haematologica 2003; 88(9):994-1001.
62. Meeker TC, Hardy D, Willman C et al. Activation of the interleukin-3 gene by chromosome translocation in acute lymphocytic leukemia with eosinophilia. Blood 1990; 76(2):285-289.
63. Akasaka T, Balasas T, Russell LJ et al. Five members of the CEBP transcription factor family are targeted by recurrent IGH translocations in B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Blood 2007; 109(8):3451-3461.
64. Raimondi SC, Behm FG, Roberson PK et al. Cytogenetics of childhood T-cell leukemia. Blood 1988; 72(5):1560-1566.
65. Hatano M, Roberts CW, Minden M et al. Deregulation of a homeobox gene, HOX11, by the t(10; 14) in T-cell leukemia. Science 1991; 253(5015):79-82.
66. Dube ID, Kamel-Reid S, Yuan CC et al. A novel human homeobox gene lies at the chromosome 10 breakpoint in lymphoid neoplasias with chromosomal translocation t(10; 14). Blood 1991; 78(11):2996-3003.
67. Kennedy MA, Gonzalez-Sarmiento R, Kees UR et al. HOX11, a homeobox-containing T-cell oncogene on human chromosome 10q24. Proc Natl Acad Sci USA 1991; 88(20):8900-8904.
68. McGuire EA, Hockett RD, Pollock KM et al. The t(11; 14)(p15; q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including Ttg-1, a gene encoding a potential zinc finger protein. Mol Cell Biol 1989; 9(5):2124-2132.
69. Royer-Pokora B, Loos U, Ludwig WD. TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11; 14)(p13; q11). Oncogene 1991; 6(10):1887-1893.
70. Van Vlierberghe P, van Grotel M, Beverloo HB et al. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood 2006; 108(10):3520-3529.
71. Boehm T, Foroni L, Kaneko Y et al. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. Proc Natl Acad Sci USA 1991; 88(10):4367-4371.
72. Bernard O, Barin C, Charrin C et al. Characterization of translocation t(1; 14)(p32; q11) in a T and in a B acute leukemia. Leukemia 1993; 7(10):1509-1513.
73. Aplan PD, Lombardi DP, Ginsberg AM et al. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. Science 1990; 250(4986):1426-1429.
74. Mellentin JD, Smith SD, Cleary ML. lyl-1, a novel gene altered by chromosomal translocation in T-cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. Cell 1989; 58(1):77-83.
75. Xia Y, Brown L, Yang CY et al. TAL2, a helix-loop-helix gene activated by the (7; 9)(q34; q32) translocation in human T-cell leukemia. Proc Natl Acad Sci USA 1991; 88(24):11416-11420.
76. Wang J, Jani-Sait SN, Escalon EA et al. The t(14; 21)(q11.2; q22) chromosomal translocation associated with T-cell acute lymphoblastic leukemia activates the BHLHB1 gene. Proc Natl Acad Sci USA 2000; 97(7):3497-3502.
77. Baer R. TAL1, TAL2 and LYL1: a family of basic helix-loop-helix proteins implicated in T-cell acute leukaemia. Semin Cancer Biol 1993; 4(6):341-347.
78. O'Neil J, Billa M, Oikemus S et al. The DNA binding activity of TAL-1 is not required to induce leukemia/lymphoma in mice. Oncogene 2001; 20(29):3897-3905.
79. Bain G, Engel I, Robanus Maandag EC et al. E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas. Mol Cell Biol 1997; 17(8):4782-4791.
80. Yan W, Young AZ, Soares VC et al. High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice. Mol Cell Biol 1997; 17(12):7317-7327.
81. O'Neil J, Shank J, Cusson N et al. TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB. Cancer Cell 2004; 5(6):587-596.
82. Hawley RG, Fong AZ, Lu M et al. The HOX11 homeobox-containing gene of human leukemia immortalizes murine hematopoietic precursors. Oncogene 1994; 9(1):1-12.
83. Hawley RG, Fong AZ, Reis MD et al. Transforming function of the HOX11/TCL3 homeobox gene. Cancer Res 1997; 57(2):337-345.

84. Kawabe T, Muslin AJ, Korsmeyer SJ. HOX11 interacts with protein phosphatases PP2A and PP1 and disrupts a G2/M cell-cycle checkpoint. Nature 1997; 385(6615):454-458.
85. Riz I, Hawley RG. G1/S transcriptional networks modulated by the HOX11/TLX1 oncogene of T-cell acute lymphoblastic leukemia. Oncogene 2005; 24(36):5561-5575.
86. Bernard OA, Busson-LeConiat M, Ballerini P et al. A new recurrent and specific cryptic translocation, t(5; 14)(q35; q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia. Leukemia 2001; 15(10):1495-1504.
87. MacLeod RA, Nagel S, Kaufmann M et al. Activation of HOX11L2 by juxtaposition with 3'-BCL11B in an acute lymphoblastic leukemia cell line (HPB-ALL) with t(5; 14)(q35; q32.2). Genes Chromosomes Cancer 2003; 37(1):84-91.
88. Hansen-Hagge TE, Schafer M, Kiyoi H et al. Disruption of the RanBP17/Hox11L2 region by recombination with the TCRdelta locus in acute lymphoblastic leukemias with t(5; 14)(q34; q11). Leukemia 2002; 16(11):2205-2212.
89. Ferrando AA, Neuberg DS, Staunton J et al. Gene expression signatures define novel oncogenic pathways in T-cell acute lymphoblastic leukemia. Cancer Cell 2002; 1(1):75-87.
90. Soulier J, Clappier E, Cayuela JM et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood 2005; 106(1):274-286.
91. Speleman F, Cauwelier B, Dastugue N et al. A new recurrent inversion, inv(7)(p15q34), leads to transcriptional activation of HOXA10 and HOXA11 in a subset of T-cell acute lymphoblastic leukemias. Leukemia 2005; 19(3):358-366.
92. Ohnishi H, Kawamura M, Ida K et al. Homozygous deletions of p16/MTS1 gene are frequent but mutations are infrequent in childhood T-cell acute lymphoblastic leukemia. Blood 1995; 86(4):1269-1275.
93. Cayuela JM, Madani A, Sanhes L et al. Multiple tumor-suppressor gene 1 inactivation is the most frequent genetic alteration in T-cell acute lymphoblastic leukemia. Blood 1996; 87(6):2180-2186.
94. Ellisen LW, Bird J, West DC et al. TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. Cell 1991; 66(4):649-661.
95. Burnett RC, Thirman MJ, Rowley JD et al. Molecular analysis of the T-cell acute lymphoblastic leukemia-associated t(1; 7)(p34; q34) that fuses LCK and TCRB. Blood 1994; 84(4):1232-1236.
96. Burnett RC, David JC, Harden AM et al. The LCK gene is involved in the t(1; 7)(p34; q34) in the T-cell acute lymphoblastic leukemia derived cell line, HSB-2. Genes Chromosomes Cancer 1991; 3(6):461-467.
97. Asnafi V, Beldjord K, Libura M et al. Age-related phenotypic and oncogenic differences in T-cell acute lymphoblastic leukemias may reflect thymic atrophy. Blood 2004; 104(13):4173-4180.
98. Larson RC, Osada H, Larson TA et al. The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice. Oncogene 1995; 11(5):853-862.
99. Aplan PD, Jones CA, Chervinsky DS et al. An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice. EMBO J 1997; 16(9):2408-2419.
100. Carroll AJ, Crist WM, Link MP et al. The t(1; 14)(p34; q11) is nonrandom and restricted to T-cell acute lymphoblastic leukemia: a Pediatric Oncology Group study. Blood 1990; 76(6):1220-1224.
101. Tycko B, Smith SD, Sklar J. Chromosomal translocations joining LCK and TCRB loci in human T-cell leukemia. J Exp Med 1991; 174(4):867-873.
102. Le Coniat M, Della Valle V, Marynen P et al. A new breakpoint, telomeric to TEL/ETV6, on the short arm of chromosome 12 in T-cell acute lymphoblastic leukemia. Leukemia 1997; 11(8):1360-1363.
103. Karrman K, Andersson A, Bjorgvinsdottir H et al. Deregulation of cyclin D2 by juxtaposition with T-cell receptor alpha/delta locus in t(12; 14)(p13; q11)-positive childhood T-cell acute lymphoblastic leukemia. Eur J Haematol 2006; 77(1):27-34.

Index

A

Accessibility 4, 33-35, 59, 60, 62, 63, 65, 68, 73, 75-77, 82, 84, 90, 93-99, 103, 105-113, 116, 121, 123, 124, 126-128, 133-144, 149, 150, 153, 159, 160, 166, 168, 172, 173, 176, 181, 188
Accessibility to recombination 76, 135, 144
Acute lymphoblastic leukemia (ALL) 180, 182-187, 189, 190
Allele 34, 38, 40, 53, 60, 63, 64, 66-69, 74, 75, 94, 104, 105, 109, 111-113, 118-123, 125-127, 133, 134, 136, 137, 139-144, 159, 160, 172-174, 183, 184, 189
Allelic exclusion (AE) 34, 68, 104, 105, 111-113, 116, 117, 121, 125-128, 134, 136, 143, 144, 157, 160, 162, 166, 168, 171, 172, 176
Alternative NHEJ *see* Nonhomologous end joining
Antibody 60, 68, 73-76, 79, 83, 168, 171, 174
Antigen receptor gene evolution 174
Antisense transcription 64, 66, 69, 137, 138, 140
Artemis 2, 3, 17, 18, 26, 27, 36, 37, 46, 48, 49, 51, 52, 170
ATM 17, 34, 36, 39, 40, 52
Atypical SCID 49

B

Base-flipping 6-9, 11
Basic helix-loop-helix (bHLH) 148, 187, 189
B-cell 24, 26, 33-35, 40, 46-50, 54, 60, 62-64, 66-69, 74-77, 79, 94, 95, 104, 105, 108, 109, 111-113, 116, 122, 126, 133, 134, 136-144, 148-153, 159, 160, 168, 169, 171-174, 180, 182-186
  development 48, 60, 77, 105, 112, 133, 134, 136-139, 141-143, 150, 160, 169, 186
  fate commitment 138, 139
  fate specification 133, 134, 137, 138, 140

C

Cancer 11, 49, 53, 54, 91, 128, 180, 181, 188
Cell cycle 36, 38, 40, 52, 127, 134, 141, 142, 157, 158
Cernunnos 2, 3, 17, 37, 46, 49, 51, 53, 54
Cernunnos/XLF 53, 54
c-Fos 10
Chemokine 134, 142, 143
Chromatin 33-36, 38, 40, 52, 59, 62, 63, 65-69, 73, 75, 76, 78-80, 90, 93-99, 103, 106-113, 121-124, 126-128, 133-141, 144, 149, 153, 159-162, 172-174, 181, 184, 186, 188
  accessibility 63, 68, 90, 93, 106, 107, 109-112, 121, 123, 124, 126-128, 135, 141, 149, 153, 181, 188
  immunoprecipitation (ChIP) 73, 76, 78, 79, 122, 124, 128, 152
  remodeling 59, 62, 63, 65, 69, 78, 80, 95, 99, 109, 110, 113, 123, 124, 149
  structure 35, 40, 62, 65, 67, 75, 76, 78, 79, 93, 95-98, 123, 124, 133-140, 144, 160
Chromosomal translocation 16, 19, 21-23, 32, 54, 103, 189
Chromosome territory 38-40, 63
cis-Regulatory element 119, 122-125, 128, 144, 150
Cryptic RSS 34-36, 186
CXCR4 134, 143
CyclinA/Cdk2 158

D

Disintegration 19, 21, 25
DNA 1-6, 8-11, 16-27, 32-40, 46, 47, 49-54, 59, 60, 62, 68, 69, 74, 75, 77, 79, 80, 82, 84-86, 88, 89, 95, 96, 99, 103, 104, 107-109, 111, 116, 119, 123-128, 135, 136, 140, 142, 144, 148, 149, 152, 157-160, 162, 167-172, 174, 175, 181, 183, 185-188
  coding 18, 24, 26
  damage sensing 34, 39

footprinting 142, 149
ligase IV 2, 3, 18, 27, 36, 37
looping 68, 69, 125, 136, 140
methylation 159
repair 19, 32, 34, 37, 39, 46, 47, 50-53, 104, 157, 158, 168, 188
DNA-PK 36, 38, 51, 52, 54
DNA-PKc 2, 3, 36, 37, 38, 51, 52
D-region 66, 67

E

E2A 77-79, 94, 134, 137, 140, 142, 143, 148-153, 186, 188, 189
E12 148, 149
E47 148-150
Eα 94, 97
EBF 77, 78, 134, 137-141, 151, 153
E-box site 148-150, 153
Electrophoretic mobility shift assay (EMSA) 3
Enhancer 22, 35, 62, 63, 66, 68, 69, 79, 90, 93, 94, 97, 98, 106, 109-112, 117-119, 121-123, 126, 134, 136, 137, 140-144, 148-150, 153, 159, 172, 175, 188
Epigenetic control 73, 160
Epigenetic regulation 108, 111, 159
E-protein 148-153
Ezh2 68, 69, 77, 134, 140

F

Fluorescence in situ hybridization (FISH) 63, 64, 109, 111, 112, 126, 141, 144, 188, 189
Fluorescence resonance energy transfer (FRET) 5, 7

G

γδT cell 126, 152, 166, 168, 182, 184, 189, 190
Gene targeting 66, 94, 116, 117, 119, 120, 122, 128, 137
Genomic instability 16, 19, 33, 37, 53, 59, 186
Germline transcription 62, 63, 65, 66, 68, 69, 76, 93-95, 97, 105, 106, 109-111, 123, 124, 127, 135, 139, 142-144, 149, 160
GTP 25

H

H3K4me3 27, 108, 158, 160-162
H3R2me2s/K4me3 27
Haemophilus influenzae type b (*Hib*) 74, 75
Heavy chain exclusion 169
HEB 122, 148, 149-153, 188
Histone acetylation 66, 68, 69, 76-78, 96, 97-99, 137-139, 142, 143, 160
Histone methylation 96, 98, 161
Histone modification 35, 62, 63, 65, 66, 69, 77, 80, 96, 99, 107-109, 112, 135, 137, 139, 160, 161
HMGB1 1-8, 10, 11, 33
HMGB2 1
Homologous recombination 38, 94, 97, 158
Human primary immune deficiency 47

I

Id protein 148
IgH enhancer 137, 150
Igk enhancer 149, 153
Ig/TCR 123, 126, 171, 180-182, 184-190
IL-7R signaling 138, 140, 141
Immunoglobulin (Ig) 1, 16, 22, 32-34, 39, 46, 48, 59, 60, 68, 73-76, 98, 103-106, 108, 109, 111, 116, 117, 123, 126, 127, 133-136, 140-144, 148-150, 152, 153, 159-161, 166-176, 180-182, 184-190, 196
Immunoglobulin (Ig) gene 75, 104, 105, 109, 133, 135, 136, 144, 167, 168, 170, 171, 174, 175, 187
Immunoglobulin heavy chain (Igh) 39, 59-69, 94, 99, 116, 122, 126, 140, 148, 161, 169
Inner nuclear membrane (INM) 109, 135, 136, 140, 141, 144
Intergenic transcription 62-66, 68, 69, 94, 137
IRF4 134, 136, 142, 143, 153
IRF8 142
Iμ 'supergene' 63, 66, 68

J

Jα segment 97, 98, 107

K

Ku 2, 3, 17, 18, 27, 36-39, 51

L

Leukemia 32, 33, 35, 39, 103, 180, 184, 185, 189
Locus 10, 19, 26, 35-37, 47, 51, 59-64, 66-69, 73, 75, 77-79, 82-90, 93-95, 97-99, 103-109, 111-113, 116-120, 122-128, 137-144, 149-151, 153, 157, 159, 161, 162, 169, 171-174, 182, 186, 188, 189
  accessibility 35, 59, 60, 93, 106, 124, 142-144
  compaction 140
  contraction 35, 67-69, 79, 111, 113, 126, 128, 139-141, 169, 172-174
Lymphocyte 1, 10, 24, 32, 35, 39, 40, 46-51, 53, 54, 59, 60, 67, 68, 76, 82, 83, 85, 90, 103-106, 109, 111, 116, 117, 122, 126, 133, 144, 148, 149, 153, 157, 166, 168, 171-173, 176, 184
Lymphocyte development 1, 10, 32, 76, 103, 105, 109, 116, 133, 144, 148, 149
Lymphoid cancer 180, 181, 188
Lymphoid clonality 180, 181, 183, 184, 190
Lymphoid clonality analysis 180, 181, 183, 190
Lymphoma 26, 32, 33, 37, 39, 40, 54, 103, 180, 185, 188, 189, 190

M

Mechanism 1, 3, 5-7, 9, 11, 16, 24, 25, 27, 32-37, 39, 46, 59, 60, 62-64, 68, 69, 76, 79, 83, 90, 93, 95-99, 103-105, 107, 109-113, 116, 119, 123, 126, 128, 133-135, 137-144, 150, 151, 153, 157, 158, 162, 168, 171, 173, 174, 176, 186, 188, 190
Microhomology 37, 38
Mutagenesis 4, 8, 16, 19, 21, 24, 25, 52, 117

N

Noncoding RNA transcription 61-63, 69
Non-Hodgkin's lymphoma (NHL) 180, 185, 188
Nonhomologous end joining (NHEJ) 2, 3, 16-20, 22, 26, 27, 32-34, 36-38, 46, 51, 53, 54, 116, 158, 170
  alternative 32, 37, 38
  deficiency 37
Nuclear compartmentalization 109, 133, 134-136, 144
Nuclear lamina 135, 137, 139-141
Nuclear organization 67, 138
Nucleosome 4, 35, 62, 66, 76, 79, 95-97, 99, 107, 108, 110, 124, 134, 135, 138, 139, 153, 159, 160
Nucleosome phasing 159, 160

O

Oct-1 138, 139
Omenn syndrome (OS) 48-50
Oncogenesis 16, 19, 22, 24, 32-37, 117, 180, 181, 184-186, 188-190

P

Pathway 2, 3, 16, 18-24, 26, 27, 32, 34, 36-39, 46, 47, 122, 127, 133, 134, 138, 139, 142-144, 167, 168, 170
Pax5 10, 35, 67-69, 77, 79, 80, 94, 99, 134, 135, 139-142, 151, 153, 160
PDb1 95, 99
Pericentromeric heterochromatin 69, 112, 127, 135, 136, 142, 144
Peripheral T-lymphocyte 82, 85
PHD finger 35, 96, 99, 161, 162
Plant homeodomain 27, 35, 108, 158, 161
Post-cleavage complex 9
Pre-B-cell checkpoint 141
Pre-BCR signaling 141-143
Pre-Tα 150, 151
Pre-Ta expression 151
Promoter 22, 32, 35, 60, 62-65, 79, 80, 90, 93-99, 106-111, 117-119, 121-123, 127, 135, 138-140, 149, 153, 159, 188, 189
Proteasomal degradation 158
Protein-DNA complex 3, 38

R

Rag1 and Rag2 leaky mutation 48
Rearrangement 10, 11, 24, 32-36, 40, 46-49, 69, 73-80, 82-90, 94, 95, 97, 98, 104, 105, 107, 109-112, 116-119, 121, 122, 125-127, 133, 134, 137-144, 148-153, 157, 159, 160, 166, 168-176, 180-186, 188-190
Recombinase 6, 16, 17, 33-37, 60, 63, 68, 79, 93, 94, 97, 99, 103-109, 111-113, 116, 117, 119, 123, 128, 134, 135, 137-139, 141, 143, 149, 157, 159-162, 166, 167, 176, 180, 181, 184-190
Recombination 1-6, 8-11, 16-20, 22, 24, 27, 32-40, 46-48, 50-54, 59-69, 73-79, 82-85, 87, 90, 93-95, 97-99, 103-113, 116, 117, 119, 121-128, 133-144, 148-153, 157-162, 166-176, 180, 186
Recombination activating gene (RAG) 1-11, 16-27, 32-40, 48, 59, 60, 76, 79, 80, 94, 95, 99, 104-106, 108, 113, 123, 134, 135, 137, 140-143, 148-151, 157-162, 166-168, 170-174, 176, 186, 188, 190
  expression 149, 151
  transposon 167
Recombination activating gene 1 (RAG-1) 1, 3-9, 11, 16-21, 24-27, 33, 35, 37-39, 46-51, 60, 77, 79, 104, 116, 117, 122, 123, 133-135, 139, 149, 150, 157, 159, 167, 168, 186, 188-190
Recombination activating gene 2 (RAG-2) 1, 3-5, 7, 9, 10, 16, 17, 19, 24, 25, 27, 33, 35-39, 46-51, 60, 77, 79, 99, 104, 108, 116, 127, 128, 133-135, 139, 149, 150, 157-159, 161, 162, 167, 168
  C-terminus 24, 25, 27, 99
  deficiency 39, 47, 48
Recombination signal sequence (RSS) 1-11, 16-21, 23, 24, 26, 33-36, 46, 59, 60, 73-76, 78-80, 93, 94, 95, 98, 99, 104, 105, 108, 110, 111, 116-121, 124, 126, 128, 134, 135, 137-140, 149, 157, 159, 160, 166-170, 175, 186, 188
Recombination substrate 9, 51, 74, 75, 94
Regulation 16, 24, 25, 27, 32-34, 40, 52, 59, 60, 62, 64, 85, 90, 93, 97, 98, 103, 104, 107-109, 111-113, 116, 117, 123, 126, 128, 136, 141-143, 148-153, 157-159, 161, 162, 168, 189
RNA-FISH 63, 64
RNA polymerase II 63
RORγt 151, 153
RS-SCID 47, 49, 51-54

S

SCID with microcephaly 49
Secondary IgL rearrangement 150
Severe combined immune deficiency (SCID) 37, 46-54
Size and diversity T cell receptor repertoire 90
Skp2-SCF 158
Stat5 94, 139, 141, 143
Surrogate light chain 104, 134, 150, 151
Synaptic complex 1-7, 9-11, 33, 36

T

Target capture 19, 20, 24-27
Target site 19, 21, 24, 25, 27, 35, 158
T-B-SCID 47-50
T-cell receptor (TCR) 1, 10, 16, 32-34, 39, 46, 48-50, 59, 60, 74-76, 82-85, 87, 90, 98, 103-106, 108, 109, 111, 116, 118, 126, 127, 148-153, 166, 168, 171, 172, 180-190
  gene 16, 32, 75, 90, 103-106, 149, 150, 153, 168, 187
  loci 34, 50, 60, 74, 76, 98, 106, 108, 109, 148-150, 152, 153, 181, 185, 186, 189
  TCRα 104, 106, 107, 111, 149-151, 153
  TCRβ 10, 34, 36, 76, 83, 94, 95, 104-106, 109-113, 117, 119-121, 126, 127, 149-151, 159, 160, 172, 182, 184, 185, 188-190
Tcra 60, 66, 69, 97, 98, 117, 121, 126
TCRAD locus 82, 83-86, 88, 89
Tcrb 60, 69, 94, 95, 99, 116-119, 121-128
TdT expression 150
T early α (TEA) 90, 97-99
Thymocyte 34, 48-50, 94, 95, 104, 108-112, 116, 119, 123-127, 151, 153, 158, 174
Transcription 10, 11, 33, 35, 61-69, 73, 76-80, 84, 90, 93-99, 105-107, 109-113, 117, 119-124, 127, 133-144, 148-151, 159, 160, 162, 174

Transcriptional elongation 96-99
Transcriptional interference 98, 128
Transcriptional regulation 153
Transcription factor (TF) 10, 11, 35, 62, 67, 73, 77-80, 94, 96, 110, 119, 122, 123, 127, 133-144, 148, 151, 160
Transcription factory 68
Transcription terminator 97, 98
Transposition 8, 11, 16, 18-27
Transposon 6, 24, 27, 167

V

Variation 33, 73, 74, 79, 80, 87, 117, 159, 171
V(D)J 16-19, 22, 24, 26, 27, 32-40, 46-54, 59-62, 64-67, 69, 75, 79, 83, 93, 94, 97, 99, 103-109, 111-113, 116, 117, 119, 122-125, 126, 128, 134, 135, 138, 144, 148-153, 157-162, 166-168, 170-173, 176, 180, 181, 185, 186, 188, 190
V(D)J rearrangement 2, 10, 11, 75, 122, 172, 173, 176, 181, 185, 188
V(D)J recombination 1-6, 8-11, 16-19, 22, 24, 27, 32-40, 46-48, 50-54, 59-62, 64-67, 69, 79, 83, 93, 94, 97, 99, 103-109, 111-113, 116, 117, 123-126, 128, 134, 144, 148-153, 157-162, 166, 168, 170, 171, 173, 176, 180
V_H gene 10, 60, 62-65, 67-69, 73-75, 77, 79, 94, 112, 134-141, 169, 172, 173
$V\kappa$ gene 94, 111
V region 60, 63, 64, 66-69, 82, 84, 166, 171

X

XRCC4 2, 3, 17, 18, 26, 27, 36-38, 51, 53, 54

Y

YY1 67-69, 79, 134, 136, 140, 141