J Mol Evol (1995) 41:556–562

JOURNAL OF MOLECULAR EVOLUTION

Concerted Evolution in Protists: Recent Homogenization of a Polyubiquitin Gene in *Trichomonas vaginalis*

Patrick J. Keeling, W. Ford Doolittle

Canadian Institute for Advanced Research, Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 4H7

Received: 26 February 1995 / Accepted: 12 May 1995

Abstract. Ubiquitin is a 76-amino-acid protein with a remarkably high degree of conservation between all known sequences. Ubiquitin genes are almost always multicopy in eukaryotes, and often are found as polyubiquitin genes—fused tandem repeats which are coexpressed. Seventeen ubiquitin sequences from the amitochondrial protist *Trichomonas vaginalis* have been examined here, including an 11-repeat fragment of a polyubiquitin gene. These sequences reveal a number of interesting features that are not seen in other eukaryotes. The predicted amino acid sequences lack several universally conserved residues, and individual units do not always encode identical peptides as is usually the case. On the nucleotide level, these repeats are in general highly variable, but one region in the polyubiquitin is extremely homogeneous, with seven repeats absolutely identical. Such extended stretches of homogeneity have never been observed in ubiquitin genes and since substitutions are common in other coding units, it is likely that these repeats are the product of a very recent homogenization or amplification.

Key words: Protests — *Trichomonas vaginalis* — Polyubiquitin gene

Introduction

Eukaryotic genomes contain numerous families of repeated sequences including nongenic repeats, selfish elements, noncoding structural elements, and some multicopy genes. In many instances the duplication events that led to the amplification of these repetitive sequences took place before the divergence of species that now harbor them. Nevertheless, the individual units contained within one species will usually be more similar to one another than they are to any related sequence in any other species. This has been demonstrated in a variety of cases, including rDNA operons, multicopy protein coding genes, centromeres, and other repetitive sequences (Coen et al. 1982; Michelson and Orkin 1983; Gonzalez et al. 1989; Warburton et al. 1993). Such sequences are said to evolve in concert, diverging as a group rather than independently. The mechanisms thought to be responsible for this are unequal crossing over, gene conversion, and transposition (Dover 1982). Unequal crossover events will usually affect tandem repeats and lead to an allelic heterogeneity in the number of repeats. Gene conversion can also operate on nonallelic repeats and can homogenize sequences without necessarily altering their frequency in the genome. Transposition (or episodic pseudogene formation) generally leads to changes in the number of repeats by the creation of new copies.

Ubiquitin is a small, highly conserved protein, originally named for its presence in all cell types (Schlesinger and Goldstein 1975; Goldstein et al. 1975). For many years this name appeared to be an exaggeration as ubiquitin and its associated pathways could not be detected in prokaryotes until recently (Wolf et al. 1993; Durner and Böger 1995). Ubiquitin has many roles in the eukaryotic cell: some are general (Murti et al. 1988; Davie and Murphy 1990; Ghislain et al. 1993; Sommer and Jentsch 1993; Holloway et al. 1993), while others are specific to certain cell types, and must have been assumed only recently (Ball et al. 1987; Kelly et al. 1991; Früh et al.

Correspondence to: P.J. Keeling

1994). The growing body of evidence that ubiquitin plays a role in the 20S proteasome-mediated proteolytic pathway in the archaebacterium *Thermoplasma acidophilum* indicates that this function, at least, is ancestral to all eukaryotes (Wenzel and Baumeister 1993).

Eukaryotic ubiquitin genes are found in three major forms: as isolated open reading frames, as fusions with genes for small ribosomal proteins, or as polymers of head-to-tail ubiquitin coding sequences which are cotranslated and converted to monomers by proteolysis. In most eukaryotes there are many ubiquitin genes of one sort or another (the exception being *Giardia;* Krebber et al. 1994), and these are evolving in concert. Since ubiquitin repeats are highly conserved, relatively short (228 nucleotides), often found in tandem, and have been sequenced from many taxa, these genes are an ideal model in which to study concerted evolution (Sharp and Li 1987).

In this report, the sequences of several *Trichomonas vaginalis* ubiquitin genes are shown to be unusual in a number of interesting ways. First, the predicted polypeptide sequence is highly divergent, with several unique substitutions at positions of known function. Second, the polypeptide sequences of the monomers are not identical, as is the case in other species with very few exceptions (Neves et al. 1990; Müller et al. 1994). Last, the nucleotide sequences of portions of the polyubiquitin are absolutely conserved, while others are highly diverged from one another.

Results

Amplification of Trichomonas *ubiquitin*

Under the standard reaction conditions using the forward and reverse primers, UbA and Ub3, a single amplification product was observed corresponding to a single ubiquitin coding region (Fig. 1). This product was cloned and hybridized to a Southern blot of genomic *Trichomonas vaginalis* DNA to confirm its provenance (data not shown), and seven independent copies were sequenced. The nucleotide sequence of those clones varies at 35 out of 121 positions, resulting in three distinct polypeptide sequences. Variability in amino acid sequence is unusual in ubiquitin genes, but the nucleotide sequences were still very similar to one another, which argues that they all exist together in *Trichomonas* as part of a multicopy family. The possible presence of a polyubiquitin was addressed by amplifying under the same conditions with a set of primers, rub1 and rub2, designed to detect fused genes. Three products of the expected size were isolated and sequenced. The sequence confirmed the presence of at least one polyubiquitin locus as each contained the 3′ end of one ubiquitin gene fused to the 5′ end of another.

To obtain a minimal estimate of the number of repeats

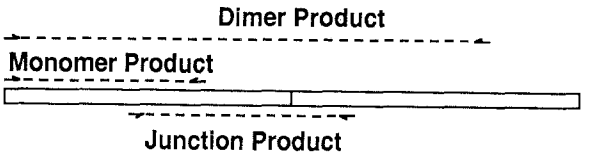

Fig. 1. Three types of amplification products characterized: Monomer products correspond to 156 nucleotides of a single ubiquitin repeat amplified using UbA and Ub3; dimer products correspond to a full-length repeat fused 156 nucleotides from the 5′ end of a second repeat obtained with the same primers; junction products are amplified using the primers Rub1 and Rub2 resulting in 118 nucleotides from the 3′ end of one repeat fused to 38 nucleotides from the 5′ end of the downstream repeat. Product sizes are somewhat larger than these sizes in the first two cases, as the primers UbA and Ub3 also have 5′ terminal restriction sites.

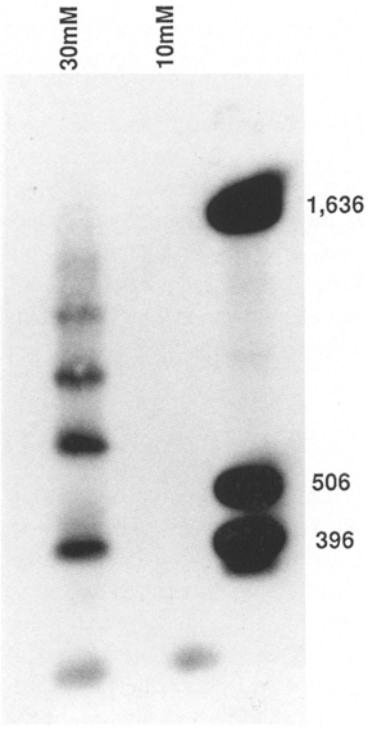

Fig. 2. Amplification of polyubiquitin. Southern blot of PCR reactions hybridized to a cloned monomeric ubiquitin gene from *T. vaginalis.* This shows the effects of the altered reaction buffer compared to the standard reaction buffer (lanes 1 and 2, respectively). The reaction carried out with 30 mM Tris (see Materials and Methods) yields products ranging from a single monomer (173 nucleotides) to at least six tandem repeats (1,313 nucleotides), whereas the reaction performed in the standard 10 mM Tris yields only a single product corresponding to a single ubiquitin gene.

in the locus, the amplification reaction buffer was optimized for amplification of products composed of more than a single unit. As indicated here, the standard reaction conditions result in a single, monomer-sized product, but we found that by increasing the amount of Tris buffer in the reaction, larger multimers could be preferentially amplified. Figure 2 shows that these conditions result in products ranging in size from a single ubiquitin unit to at least six head-to-tail repeats. This not only

Table 1. Redundancy of amplification products

Monomer
1A
1C
1D = 1B
1E = 1F = 1G
Dimer
2A = 2B
2C = 2D = c9 – 10
Junction
jA = 2B
jB = jC

confirms the presence of a polyubiquitin but also gives a lower limit to its size. Several isolates from the band corresponding to a di-ubiquitin product were cloned and sequenced, and these were found to be of the expected structure.

In total, 14 amplification products were sequenced, of three different forms. Of these seven are monomeric (1A–G), four are dimeric (2A–D), and three correspond to junctions (jA–C). Of these sequences, several were represented more than once or were subsequently found to match part of the cDNA exactly: these are not included in our analysis; however, a summary of which products are identical to which is given in Table 1.

Isolation of cDNA

To characterize a larger number of intact repeats, in their natural order, we screened a cDNA library and isolated a single clone that hybridized to the previously sequenced amplification products. The cDNA clone was found to contain an insert of about 2.5 kb, which we estimated could contain ten repeats plus short upstream and downstream untranslated regions. Sequencing the cDNA revealed that the 5′ end is truncated 39 bp upstream of the terminal glycine of a coding unit, which is followed by ten intact units, the last of which ends in an extra phenylalanine residue, a stop codon, and 34 bp of untranslated sequence. The actual length of the gene is unknown but has to be 11 units or greater. This is comparatively large, but by no means the largest reported polyubiquitin locus (Wong et al. 1992).

The unique feature of the cDNA sequence is the high degree of conservation between coding units; in fact almost all the repeats are exactly the same on the nucleotide level. Of the ten complete units, seven have precisely the same sequence, and two of the other three share a block containing ten substitutions form this sequence (Fig. 3).

Concerted Evolution of Ubiquitin Repeats

To analyze this dramatic case of concerted evolution, pairwise distances between coding units were calculated and phylogenetic trees constructed. The data set used contained only the sequences that were unique or else present in the cDNA; in this way we hoped to better represent the range of sequences present in the genome by avoiding overrepresentation of preferentially amplified PCR products while at the same time considering the repetitive nature of the cDNA. The final set of data considered is shown in Fig. 4 where the repeated cDNA sequence is used as a standard to which the other sequences are aligned.

Ubiquitin is not a good marker for phylogeny as it is too short, saturated, and too highly conserved (Wray and DeSalle 1994). However, phylogenetic tree construction may still be useful to show that the members of a repeat family are all more closely related to one another than to any other known sequence. This is seen to be the case in Fig. 5, where all the nucleotide sequences from a particular taxon group together to the exclusion of all other sequences, a sign that they originated recently or are evolving in concert. It is noteworthy that *Trichomonas* units 1A and 1E, which are the two that differ at the amino acid level, are found to branch well within the *Trichomonas* cluster, and that the substitutions that unite cDNA units c4 and c10 are also common to most of the amplification products.

Table 2 addresses the range of variability found between repeats in a number of species. Similar tables can be found elsewhere (Sharp and Li 1987; Tan et al. 1993), so we have concentrated on recent data and protists, while trying to give a good representation of the range of homogeneity. The mean and upper range of pairwise distances found in *Trichomonas* are higher than those of most other taxa, while not the highest, arguing that there is a relatively high degree of variability in *Trichomonas,*

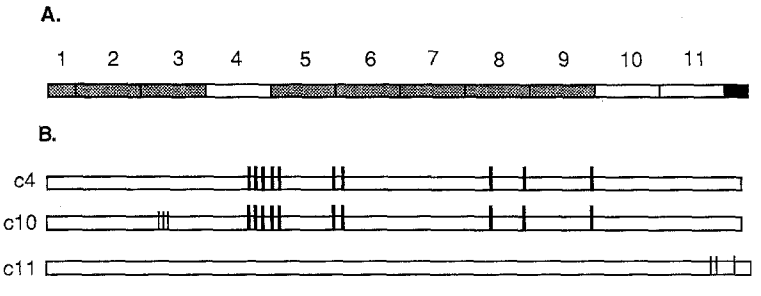

Fig. 3. Schematic diagram of the cDNA clone showing identical repeats and the position of substitutions that vary from this sequence. **A** The cDNA is composed of a 5′ truncated repeat, ten intact repeats, and a short 3′ untranslated region. The repeats shown in *gray* are all absolutely identical; the other three are shown in **B** where thin vertical marks denote a position that is unique to that repeat, and a *heavy vertical mark* denotes a position that varies from the seven identical repeats but is shared between two of the unique repeats.

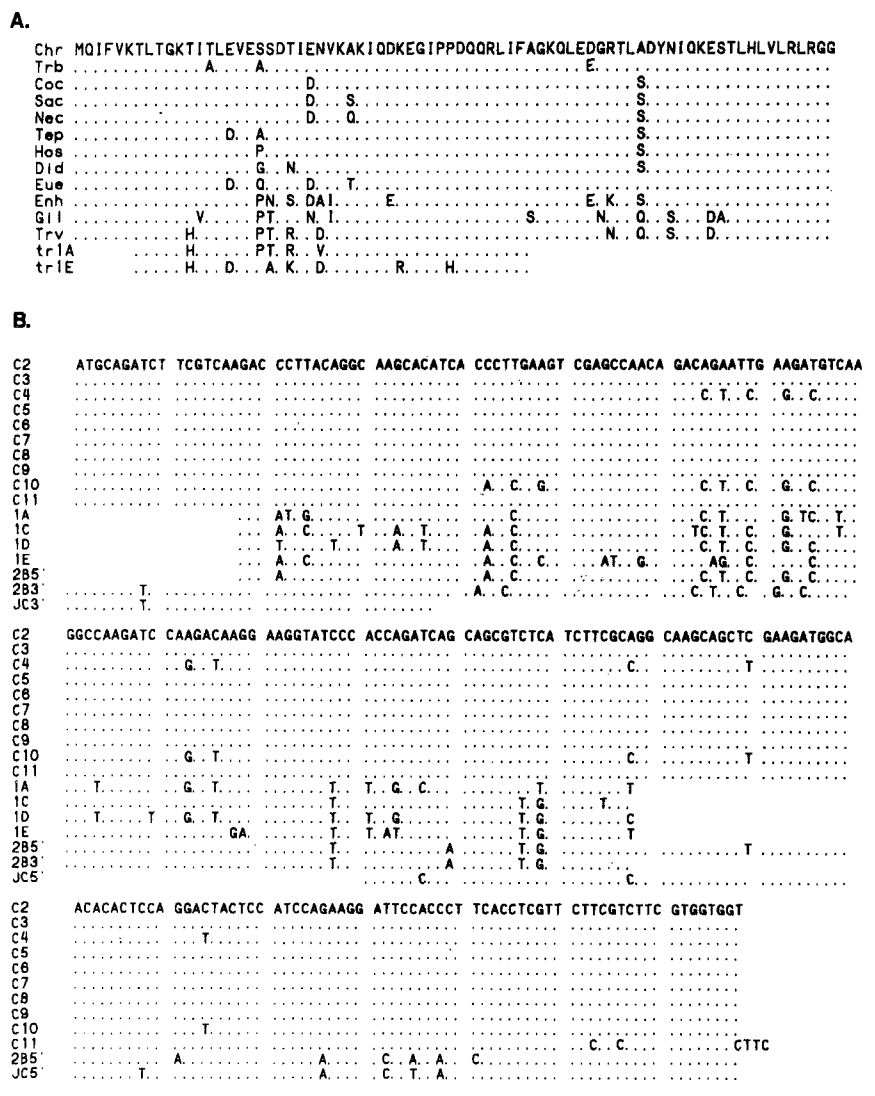

Fig. 4. Amino acid and nucleotide sequences of ubiquitin genes from *Trichomonas vaginalis*. **A** Aligned ubiquitin peptide sequences from a selection of organisms (*Chr, Chlamydomonas reinhardtii; Trb, Trypanosoma brucei; Coc, Coprinus congregatus; Sac, Saccharomyces cerevisiae; Nec, Neurospora crassa; Tep, Tetrahymena pyriformis lucos Tu2; Hos, Homo sapiens; Did, Dictyostelium discoideum; Eue, Euplotes eurystomus; Enh, Entamoeba histolytica; Gil, Giardia lamblia; Trv, Trichomonas vaginalis* cDNA; *trlA, T. vaginalis* variant 1A; *trlE, T. vaginalis* variant 1E). **B** Nucleotide sequences of *T. vaginalis* ubiquitin genes. Only unique sequences and those of the cDNA are shown.

despite the block of extremely homogenous repeats in the cDNA.

Discussion

The ubiquitin sequences from *Trichomonas vaginalis* are surprising for two reasons: in one sense they are extremely variable, while in another they are uncommonly conserved. In most polyubiquitin genes all nucleotide variation is in the form of synonymous substitutions. This is also the case in the cDNA fragment from *Trichomonas,* but several of the PCR products vary from this sequence, resulting in seven variable sites and a total of three different sequences (Fig. 4A).

The amino acid sequences themselves are interesting for there are numerous positions where the *Trichomonas* sequence contains unique substitutions, and many of these either alter the charge, or are not isosteric. One of particular interest is N54, which is an otherwise highly conserved arginine that has been demonstrated through site-directed mutagenesis to be involved in the formation of the ubiquitin-adenylate intermediate in the conjugation of ubiquitin to the ubiquitin-activating enzyme E1 (Burch and Haas 1994). The presence of an asparagine at this position in *Trichomonas* and a lysine in *Entamoeba* (Fig. 4A) raises questions about the contribution of this residue to conjugation in these organisms.

The most unusual feature of the *Trichomonas* ubiquitin genes is at the nucleotide level, specifically in the 3′ end of the polyubiquitin (Fig. 4B). Comparing the pairwise distance between repeats, it is clear that there is a high degree of interrepeat variability relative to that observed in other taxa. In light of this, the remarkable homogeneity of the cDNA, where seven out of ten repeats in this region are absolutely identical, must be considered a special case, and most likely a recent event.

Two mechanisms that drive homogenization of repeats are gene conversion and unequal crossing-over. Both of these have been described in polyubiquitin genes (Sharp and Li 1987; Baker and Board 1989), and in both cases generating a tandem array of identical repeats requires multiple events. For multiple gene conversions to

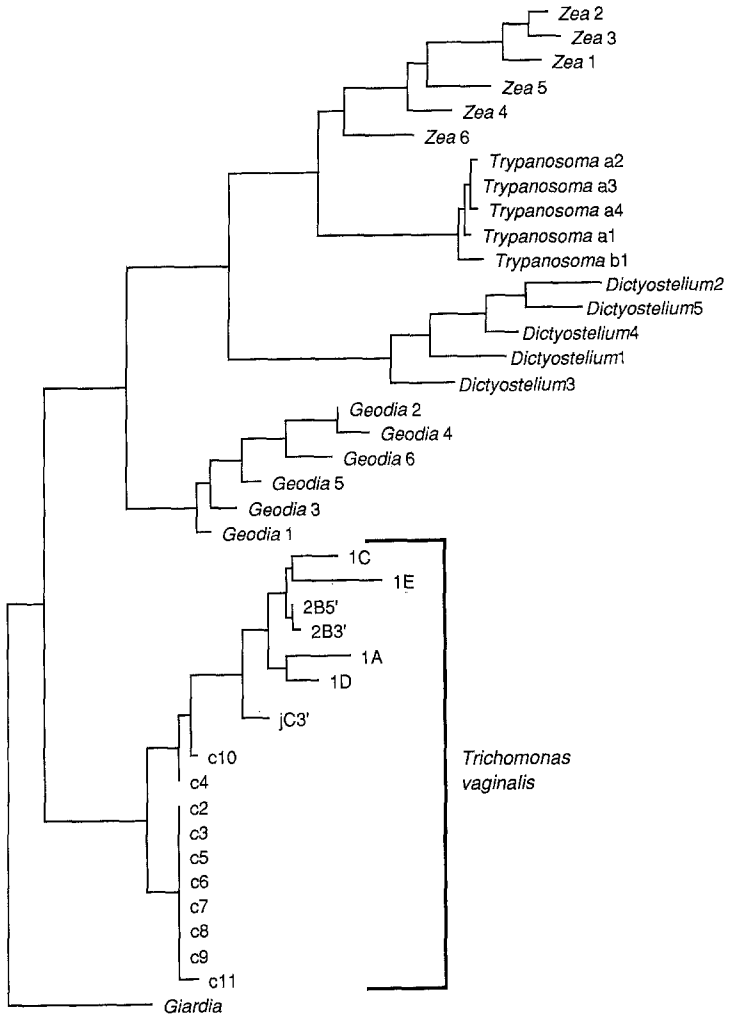

Fig. 5. Parsimony tree of ubiquitin nucleotide sequences. Polyubiquitin genes were divided into individual monomers and aligned with the nucleotide sequences from *Trichomonas;* the most parsimonious tree was found using PAUP under default conditions. The tree shows that sequences from a particular species form a coherent group (i.e., they are evolving in concert). In addition, it can be seen that variable repeats c4 and c10 from the *Trichomonas* cDNA are more similar to the amplification products than the seven identical repeats, which can be seen fairly clearly from the alignment alone.

produce the observed pattern there would also have to be a strong bias toward homogenizing repeats clustered at one location while not at others (as suggested by the high level of variation in individual PCR products compared to that of the cDNA fragment). Alternatively, unequal crossovers tend to create tandem duplications (and corresponding deletions), so several unequal crossovers in roughly the same area of the gene could easily yield tandem replications of the sort observed here. However, if the long stretches of homogeneity observed in the cDNA are apparently the product of unequal crossing-over events, the pattern of identity between cDNA repeats c4 and c10 are more likely the product of a gene conversion as two identical blocks of sequence are surrounded by somewhat different contexts. There is evidence from other repeat families that different mechanisms of homogenization can operate simultaneously (reviewed in Dover 1993). This appears also to be the case here, where there is clear evidence of both unequal crossing-over and gene conversion, but the only sure conclusion is that the evolution of this locus involved a complicated series of events. One possible evolutionary pathway is shown in Fig. 6.

In an interesting aside, the *Trichomonas vaginalis* genome is known to contain several families of repeated sequences, one of which has been partly characterized and found to be a tandem repeat that is extremely homogeneous. The 650-bp repeated units were found to vary at only 15 sites in 8 clones examined (Paces et al. 1992).

Trichomonas vaginalis is one of the earliest-diverging eukaryotic lineages (Sogin 1989; Cavalier-Smith 1993), so the presence of polyubiquitin genes in this organism shows that this unusual expression strategy is a very ancient eukaryotic feature. Ubiquitin and proteosome-like sequences are now known in the eubacteria (Lupas et al. 1994; Durner and Böger 1995), and observations of their activity in the archaebacterium, *Thermoplasma* (Wenzel and Baumeister 1993), argue that this pathway evolved *at least* by the time that eukaryotes and archaebacteria diverged, and possibly earlier. Given this, it is likely that other ubiquitin-mediated processes also evolved before eukaryotes. This cannot be said for sure, however, without more data, as important factors, such as conjugating enzymes, have only been described in higher eukaryotes.

Table 2. Distances between repeats in polyubiquitin loci

Species	Mean	Corrected mean	Range	Corrected range
Zea maize	23.5	0.103	8 to 33	0.035 to 0.145
Phytophthora infestans	16.4	0.072	5 to 24	0.022 to 0.105
Geodia cydonium	3.8	0.017	1 to 7	0.004 to 0.031
Cricetulus griseus	3.7	0.016	0 to 5	0 to 0.022
Bombyx mori	4.2	0.018	1 to 6	0.004 to 0.026
Bos taurus	32.3	0.142	30 to 35	0.132 to 0.154
Euplotes eurystomus	16.3	0.071	10 to 20	0.044 to 0.088
Tetrahymena pyriformis	63.4	0.278	36 to 95	0.158 to 0.417
Trypanosoma cruzi	7.3	0.032	7 to 8	0.031 to 0.035
Trichomonas vaginalis	20.5	0.090	0 to 45	0 to 0.198

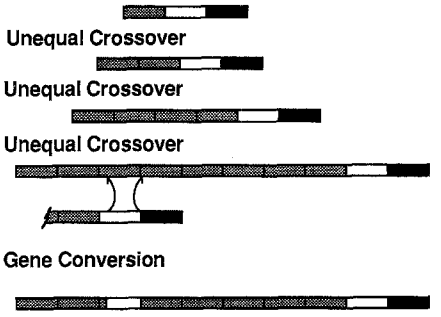

Fig. 6. Possible model for the evolution of the *T. vaginalis* polyubiquitin repeat. Repeated rounds of unequal crossing-over would result in a long tandem array of identical monomers. The actual number of these events may be greater than that shown. The similarity in sequence between repeats c4 and c10 could be generated by a gene conversion even either between these two repeats as shown, or between some other upstream repeat.

Materials and Methods

General Molecular Techniques. E. coli strain TG1 was used for all cloning and sequencing and was grown under standard conditions. Enzymes were obtained from NEB, BRL, Boeringher-Mannheim, and Appligene, and all were used according to the manufacturers' directions except where indicated. Southern transfers and plaque lifts were carried out to the manufacturer's specifications (NEN). Blots were hybridized at 65°C in 10% dextran sulfate, 1 M NaCl, 1% SDS, and washed twice at room temperature in 2× SSC (0.15 M NaCl, 0.015 M sodium citrate), twice at 65°C in 1% SDS, 2× SSC, and twice in 0.1× SSC at room temperature. Enzymatic sequencing of double-stranded templates was performed using Sequenase (USB) according to specifications provided. All manipulations of the *Trichomonas* cDNA library were according to Lambda ZAPII protocol (Promega).

Amplification of Ubiquitin Genes. All Amplification reactions consisted of 35 cycles of 92°C denaturing for 1 min, 55°C annealing for 2 min, and 72°C extension for 2 min, all followed by a 5-min extension at 72°C. Ubiquitin monomers were amplified using degenerate primers UbA CGGGATCCCGATGCA(A/G)AT(C/T/A)TT(C/T)GT(A/T/G/C)AA and Ub3 CGGGATCCCTC(T/C)TC(A/T/G/C)A(A/G)(T/C)TG(T/C)TT(A/T/G/C)CC, each based on absolutely conserved regions of all known ubiquitin protein sequences. Single junctions were amplified using *Trichomonas*-specific primers Rub1 G(T/C)TG(A/G)(T/C)TCGAC(T/G)TCGAG(T/G)GTG and Rub2 CA(G/A)GA(T/C)A(A/G)(A/G)GAAGGTATTCC. Multimers were amplified preferentially using UbA and Ub3 in reaction buffer containing 30 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100, and 0.2 mg/ml BSA at pH 9 (25°C). All PCR products were isolated by electrophoresis, purified by freezing crushed gel slices with an equal volume of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and two volumes of phenol at −70°C followed by centrifugation and ethanol precipitation of the aqueous phase. Purified products were cloned using TA cloning vector, pCRII (Invitrogen).

Terminal Deletion of Ubiquitin Repeats. 5′ terminal monomers were deleted by incubating 1 μg of cDNA clone with 15 units of *Bam*HI for 1.5 h, followed by 5 units of *Bgl*II for 20 min, at which time half the reaction was removed and stopped by the addition of 10 mM EDTA. The remainder was allowed to digest for an additional 20 min and then the two halves were pooled once again. This digest was then electrophoresed overnight in 0.7% agarose and individual bands were isolated and purified as described. The purified DNA was then circularized overnight in a dilute ligation and transformed into *E. coli.* 3′ terminal monomers were deleted in much the same way, except that the cDNA was first treated with *Bgl*II then overhangs were filled in by incubation at 37°C for 5 min with Klenow and 10 mM deoxynucleotides in the digestion reaction buffer. The DNA was then ethanol precipitated, resuspended, and digested overnight with *Hin*cII. Individual deletions isolated and ligated as described for the 5′ deletions.

Sequence Analysis. Amplification products and cDNA sequences were aligned with the nucleotide sequence of several polyubiquitin genes retrieved from GenBank. PAUP (Phylogenetic Analysis Using Parsimony; Swofford 1993) was used to calculate pairwise distances between coding units of individual taxa, and parsimony trees were constructed by a heuristic search for the shortest tree under default conditions. Since the amplified sequences did not cover the entire length of a coding unit, pairwise distance measurements on *Trichomonas* sequences were corrected for missing data to give a value per nucleotide position. Absolute distance values for *Trichomonas* given in Table 1 are projections based on the corrected distance calculation and as such are not necessarily real, but were included for comparison to other taxa.

Acknowledgments. We are grateful to P. Johnson for providing us with a *T. vaginalis* cDNA library and to M. Müller for genomic *T. vaginalis* DNA. We also thank A.L. Doherty for help with sequencing and E. Teh for technical assistance. This work was supported by a Medical Research Council (MRC) operating grant (MT4467). P.J.K. was supported by an MRC Studentship. W.F.D. is a fellow of the Canadian Institute for Advanced Research.

References

Baker RT, Board PG (1989) Unequal crossover generates variation in ubiquitin coding unit number at the human *UbC* polyubiquitin locus. Am J Hum Genet 44:534–542

Ball E, Karlik CC, Beall CJ, Saville DL, Sparrow JC, Bullard B, Fyrberg EA (1987) Arthrin, a myofibrillar protein of insect flight muscle, is an actin-ubiquitin congugate. Cell 51:221–228

Burch TJ, Haas AL (1994) Site-directed mutagenesis of Ubiquitin. Differential roles for arginine in the interaction with ubiquitin-activating enzyme. Biochemistry 33:7300–7308

Cavalier-Smith T (1993) Kingdom protozoa and its 18 phyla. Microbiol Rev 57:953–994

Coen ES, Thoday JM, Dover G (1982) Rate of turnover of structural variants in the rDNA gene family of *Drosophila melanogaster*. Nature 295:564–568

Davie JR, Murphy LC (1990) Level of ubiquitinated histone HB2 in chromatin is coupled to ongoing transcription. Biochemistry 29: 4752–4757

Dover G (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111–117

Dover G (1993) Evolution of genetic redundancy for advanced players. Curr Opin Genet Dev 3:902–910

Durner J, Böger P (1995) Ubiquitin in the prokaryote Anabaena variabilis. J Biol Chem 270:3720–3725

Dworkin-Rastl E, Shrutkowski A, Dworkin MB (1984) Multiple ubiquitin mRNAs during *Xenopus laevis* development contain tandem repeats of the 76 amino acid coding sequence. Cell 39:321–325

Früh K, Gossen M, Wang K, Bujard H, Peterson PA, Yang Y (1994) Displacement of housekeeping proteasome subunits by MHC-encoded LMPs: a newly discovered mechanism for modulating the multicatalytic proteinase complex. EMBO J 13:3236–3244

Ghislain M, Udvardy A, Mann C (1993) *S. cerevisiae* 26S protease mutants arrest cell division in G2/metaphase. Nature 366:358–362

Goldstein G, Scheid M, Hammerling U, Boyse EA, Schlesinger DH, Niall HD (1975) Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proc Natl Acad Sci USA 72:11–15

Gonzalez IL, Petersen R, Sylvester JE (1989) Independent insertion of Alu elements in the human ribosomal spacer and their concerted evolution. Mol Biol Evol 6:413–423

Holloway SL, Glotzer M, King RW, Murray AW (1993) Anaphase is initiated by proteolysis rather than by the inactivation of a maturation-promoting factor. Cell 73:1393–1402

Kelly A, Powis SH, Glynne R, Radley E. Beck S, Trowsdale J (1991) Second proteasome-related gene in the human MHC class II region. Nature 353:667–668

Krebber H, Wöstmann C, Bakker-Grunwald T (1994) Evidence for the existance of a single ubiquitin gene in *Giardia lamblia*. FEBS Lett 343:234–236

Lupas A, Zwickl P, Baumeister W (1994) Proteasome sequences in eubacteria. Trends Biochem Sci 19:533–534

Michelson AM, Orkin SH (1983) Boundaries of gene conversion within the duplicated human alpha-globin genes. Concerted evolution by segmental recombination. J Biol Chem 258:15245–15254

Müller WEG, Schröder HC, Müller IM, Gamulin V (1994) Phylogenetic relationship of ubiquitin repeats in the polyubiquitin gene from the marine sponge *Geodia cydonium*. J Mol Evol 39:369–377

Murti KG, Smith HT, Freid VA (1988) Ubiquitin is a component of the microtubule network. Proc Natl Acad Sci USA 85:3019–3023

Neves A, Guerreiro P, Rodrigues-Pousada C (1990) Striking changes in polyubiquitin genes of *Tetrahymena pyriformis*. Nucleic Acids Res 18:656

Paces J, Urbankova V, Urbanek P (1992) Cloning and characterization of a repetitive DNA sequence specific for *Trichomonas vaginalis*. Mol Biochem Parasitol 54:247–256

Schlesinger DH, Goldstein G (1975) Molecular conservation of 74 amino acid sequence of ubiquitin between cattle and man. Nature 253:423–424

Sharp PM, Li W-S (1987) Ubiquitin genes as a paradigm of concerted evolution of tandem repeats. J Mol Evol 25:58–64

Sogin ML (1989) Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. Am Zool 29:487–499

Sommer T, Jentsch S (1993) A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. Nature 365: 176–179

Swofford DL (1993) PAUP: phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign, IL

Tan Y, Bishoff ST, Riley MA (1993) Ubiquitins revisited: further examples of within- and between-locus concerted evolution. Mol Phylo Evol 2:351–360

Warburton PE, Waye JS, Willard HF (1993) Nonrandom localization of recombination events in human alpha satellite repeat unit variants: implications for higher-order structural characteristics within centromeric heterochromatin. Mol Cell Biol 13:6520–6529

Wenzel T, Baumeister W (1993) *Thermoplasma acidophilum* proteasomes degrade partially unfolded and ubiquitin-associated proteins. FEBS Lett 326:215–218

Wolf S, Lottspeich F, Baumeister W (1993) Ubiquitin in the archaebacterium *Thermopasma acidoophilum*. FEBS Lett 326:42–44

Wong S, Elgort MG, Gottesdiener K, Campbell DA (1992) Allelic polymorphism of the *Trypanosome brucei* polyubiquitin gene. Mol Biochem Parasitol 55:187–196

Wray CG, DeSalle R (1994) Phylogenetic utility of ubiquitin DNA sequences from 3 marine protist lineages. Mol Marine Biol Biotech 3:13–22