Current Genetic Medicine Reports (2018) 6:136–143
https://doi.org/10.1007/s40142-018-0154-9

CYTOGENETICS (T LIEHR, SECTION EDITOR)

Fragile Sites as Drivers of Gene and Genome Evolution

Kathleen Wilhelm[1] · Constanze Pentzold[1] · Sandra Schoener[1] · Arsen Arakelyan[2] · Anna Hakobyan[2] · Kristin Mrasek[1] · Anja Weise[1]

Published online: 25 September 2018

Abstract

Purpose of Review Although the detailed composition of the human genome is known base by base for its major part, the orchestration of and which elements exactly facilitate organization and flexibility of higher order gene and genome architecture, are poorly understood and scarcely studied.

Recent Findings This review focuses on fragile sites (FSs). They are considered as regions of chromosome breakage with overlapping signatures for breakpoints observed repeatedly in tumor and constitutional rearrangements, and also in evolutionary conserved breakpoints. Thus, FSs are promising targets to study and get deeper insights into chromosome, gene, and genome evolution.

Summary Here, we summarize the current knowledge on FSs and their correlation with aforementioned breakpoint categories. Based on that, we introduce a new model for FSs driven gene and genome evolution, which also can explain the recently observed spreading of (pseudo-)gene family members among the human genome. FSs therefore may provide an "infrastructure" to distribute gene copies onto different sites of the genome and may be the underlying cause for formation of gene families.

Keywords Fragile sites (FSs) · Chromosomal evolution · Gene evolution · Genome evolution · Breakpoint reuse

Introduction

One fundamental question that genome biology, chromosomal evolution, cancer- and aging-related research unifies, is, how large parts of genomes in the range of megabasepairs (Mb) can be evolutionary stable and at the same time easily and in parts reproducibly be (re-)organized with controversial outcomes like adverse in cancer or advantageous during speciation. Despite a huge amount of published and ongoing research concerning single aspects of this puzzle, the question how to find a balance between evolutionary advantage and fatal development towards disease and neoplasia due to genomic rearrangements is far from being understood. It is almost a truism to state that the common basis of such events is DNA double-strand breaks (DSBs) and their (im-)perfect repair. Depending on cell type—germline or somatic—such break events and their repair results either, if linked with selection benefits and/or population bottle necks, subsequently evolutionary conserved breakpoints, or a hereditable genomic variant, maybe leading to a disease. In latter case and if somatic cells are affected, the resulting rearrangement(s) may contribute to aging and/or the development of cancer.

Our own work on evolutionary conserved breakpoints [1–5], breakpoints in constitutional rearrangements [6–9] and neoplasia associated breakpoints [10, 11], as well as a series of studies from others [e.g., 12, 13, 14•, 15, 16•] revealed a high degree of overlap of so called fragile sites (FSs) and all aforementioned breakage events. FSs present as cytogenetic visible breaks and gaps, are regarded as a consequence of special features being present in chromosome biology, and are especially expressed under certain cell culture conditions. FSs are chromosomal regions containing DNA sequences, which occasionally enter mitosis before completion of replication, and are therefore "prone to break" under conditions of in vitro induced replication stress, e.g., due to

This article is part of the Topical Collection on *Cytogenetics*

✉ Anja Weise
Anja.Weise@med.uni-jena.de

[1] Institute of Human Genetics, Jena University Hospital, Am Klinikum 1, 07747 Jena, Germany

[2] Group of Bioinformatics, Institute of Molecular Biology NAS RA, 7 Hasratyan Str, 0014 Yerevan, Armenia

aphidicolin, folate-deficient media, bromodeoxyuridine (BUdR) or 5-azacytidine [17–19]. FSs appear with variant frequencies in dependence of tissue and individual person and its overall genetic background [e.g., 20•, 21, 22]. So far, more than 230 different FSs are described on a cytogenetic level, at a genomic resolution of 5–10 Mb [20•]. The main reason why only 41 of these sites are mapped by now on a molecular level is that they can be observed only in low frequencies, mostly below 0.1% of studied cells. Additionally, FSs are not linked to specific DNA sequences but are regions of enhanced breakage susceptibilities; these regions are variable for each FS and can span up to several Mb. No clinical relevance of FSs has been shown yet. Besides, these also so-called common FSs, which obviously belong to the regular chromosomal structure and biology, there are other FSs, which were aligned with certain syndromes. However, the latter are breakage prone due to trinucleotide repeats within the DNA, which can expand; they are not subject of this review and were discussed elsewhere [14•].

We speculate, that the seemingly contradictory features of on the one hand evolutionarily being advantageous but at the same time adversely leading to aging- and disease-related genome changes are two sides of the same coin: i.e., the trade-off on the narrow ridge of FSs.

Chromosomal (In)stability

The ability to keep the integrity of the genome although it is hit by thousands of external and internal DNA damage events every day is crucial for cell and organism survival. If continuous DNA repair is not successful, this can lead to chromosome instability (CIN) which is a hallmark of cancer and aging [e.g., 23]. In fact, a causative connection between chromosomal aberrations and tumorigenesis was already postulated by Theodor Boveri in 1914 [24]. CIN is also a well-known feature of human autosomal recessive inherited disorders, designated as chromosome breakage syndromes like Bloom syndrome (OMIM 210900), Ataxia teleangiectatica (OMIM 208900), or Fanconi anemia (FA) with 21 genetic subtypes (reviewed in [25•]). The later mentioned diseases all have in common an underlying DNA repair defect. Loss of function mutations in these disorders lead to a general premature aging phenotype, including a high prevalence of early onset of different cancer types in addition to an elevated irradiation and replication stress sensitivity.

Factors Contributing to FS Instability

As recently reviewed, there is no single common mechanism being responsible for the increased breakage rates of FSs (summarized in [26••]). However, several causative, contributing factors are well established, such as DNA features that impair proper DNA replication (e.g., replication timing and paucity in origins of replication), AT-rich sequences (microsatellites) or local DNA flexibility peaks. Slowing down the speed of DNA polymerases during replication is crucial, especially at high flexibility DNA stretches, which are enriched for interrupted runs of AT-dinucleotide repeats, like present at higher degree in FSs. During replication generated single-stranded, unreplicated regions can form DNA secondary structures (e.g., hairpins or cruciform) resulting in replication fork stalling [27]. Thus, proteins involved in the resolution of DNA secondary structures (e.g., WRN: helicase and exonuclease activity or BLM: helicase activity) are required for FS stability [e.g., 28–32]. Nonetheless, regarding the DNA flexibility in regions of FSs, contradictory data can be found in the literature [33, 34]. More recently, chromatin loop anchor points (LAP) as part of DNA organization in the interphase were identified to be loci of evolutionary changes, recombination hot spots in cancer and germ cells, and are discussed as a characteristic feature of a subset of FSs [35].

In addition to slow down DNA polymerases, collision between molecular machineries responsible for transcription and replication can lead to fork stalling by formation of so-called transcriptional R-loops (RNA/DNA hybrids). At this point, the FA pathway for DNA repair is activated (by FANCD2 and FANCM), which promotes DNA/RNA hybrid resolution and replication fork restart with the goal to limit R-loop accumulation [36, 37]. This is especially important for the transcription of large genes (up to 1.5 Mb, e.g., *FHIT* in FRA3B) leading subsequently to DSBs and the expression of FSs [38–40]. The impairment of DNA replication is therefore the major source of spontaneous DSBs in dividing cells. Therefore, incomplete replication can lead to deletions and translocations [41••, 42], or, because of DNA repair, copy number variants (CNVs) can appear involving these sites [26••, 43]. In agreement with this, Fungtammasan et al. [33] demonstrated in a genome-wide study that FSs are typically enriched in Alu elements. At the same time, such regions are preferentially located in CNV-associated breakpoints and can, in turn, mediate nonallelic homologous recombination (NAHR) or nonhomologous end-joining (NHEJ); the latter are also contributing to genomic diseases, cancer, and aging [14•, 44–51].

Divergent Role of FSs in Tumorigenesis and Speciation

The fragility of specific chromosomal regions in somatic cells may also lead to hazardous effects on an individual level by contributing to tumorigenesis by amplification of oncogenes; this is a point being discussed since at least 1997 [52–54]. In addition, around 50% of recurrent tumor-associated deletions originate from FS regions [55] and harbor tumor suppressor genes being involved in DNA repair, such as *FHIT* (colocalizing with FRA3B), *PARK2* (colocalizing with FRA6E), or *WWOX* (colocalizing with FRA16D) [56]. Nevertheless,

there are also reports on tumor-specific chromosome breakage that only to a small part overlaps with known FSs [57–59]. Further studies on this topic are needed to get a clearer picture whether breakage and replication stress are sources of or consequences in cancer cell development.

Chromosomal fragility and imperfect repair occurring during meiosis or shortly after fertilization can lead to fixed, and from generation to generation, inheritable changes of the genome. A correlation of FSs and inherited chromosomal aberrations was already shown [8]. Genes within or near evolutionary conserved breakpoints, specifically in the mammalian lineage, seem to have an impact on new adaptive traits, with effects on immune system, brain development or gene expression in testis [60]. Overall, this pinpoints the high potential of these regions to be changed in (evolutionary) shorter times than the surrounding (more stable) areas. Thus, they are pre-adapted to "react" faster on environmental changes and/or to respond to selective pressure.

The "fragile breakage model" introduced by Pevzner 2003 [61] predicts, that (evolutionary conserved) breakpoints are not randomly distributed but clustered in hot spots. This was proven by a series of cross-species chromosome painting fluorescence in situ hybridization (Zoo-FISH) experiments in mammalian species and in silico comparative studies [1–5, 62, 63]. The latter confirmed a striking correspondence between FS location and the position of evolutionary conserved breakpoints. Subsequently, this model was developed to an "integrative breakage model of genome architecture, reshuffling and evolution" [64]. Moreover, these breakage prone regions seem not only to be conserved in mammals [e.g., 65–66] but also beyond the mammalian lineage [67•].

Besides, another striking common feature of (primate) evolutionary conserved breakpoints [68–70], clinically associated potentially inherited rearrangements [8, 71], cancer related breakpoints [26••, 72, 73] and FSs [41••] is the enrichment of CNVs, which was attributed to imperfect repair (reviewed in [26••, 43]).

Model of Gene and Genome Evolution Driven by FSs

An analysis done here for published fine mapped FSs enabled a precise sequence-to-feature analysis. A more complex picture appeared that can reconcile the positive and negative impact of breakage prone regions in an evolutionary sense, as outlined above. Accordingly, comparing molecular genetically defined FS regions to the genome on "average regions," those containing FSs are gene poor, but at the same time accumulated disease causing, OMIM annotated genes, and show a 7.7-fold enrichment of CNVs. However, no difference in tumor associated breakpoints could be detected (Fig. 1). An enrichment of CNVs in FS regions was independently demonstrated by experimental induction of breaks and analyses by high throughput methods [41••]. This reflects the signature of DNA repair and highlights also local euchromatin duplication. Such DNA duplications may span gene-coding regions, which later can give rise to new genes and/or pseudogenes. However, the accumulation of OMIM genes has not been reported and discussed so far, but current studies on FS induction in neuronal stem cells report a preferred colocalization of FSs and genes for synaptic functioning, neural adhesion, tumor suppressor genes, and mental retardation-associated genes [74, 75], pointing towards the same direction. Given that such FS regions are prone to breakage and imperfect repair resulting in CNVs that can diverge by accumulation of variants and/or mutations, respectively, this can either contribute to adaption and positive selection, or concerning dosage-sensitive genes, to human diseases; the latter is highlighted by the enrichment in OMIM genes that we found.

The yet fine mapped FSs include more than 5000 genes and transcripts, often belonging to the same gene ontology groups. Those play a role in keratinization, epidermal development, peptide cross linking, retinoic acid signaling and regulation of cell proliferation, transcription, apoptosis, and differentiation. Particularly, gene families, single gene, and pseudogene members are located in *cis* at/nearby the same FS and/or in *trans* close to another FS in the genome. For instance, semaphorin gene family members, acting as axonal growth molecules, can be found at FRA7E (four members), FRA3H (three members), and six other FSs elsewhere in the genome (Fig. 2). We speculate that such members localized in *cis* are the first to spread (seeding or donor site), as this FS is flanked by the highest number of gene family members; but this will be subject of further bioinformatic analysis. Interestingly, the enrichment of pseudogene family members for genes located within FS regions shows an identical spreading pattern in *cis* and/or *trans* of FSs. This can be exemplified for the voltage-dependent anion channels *VDAC1* in 5q31 at FRA5C and *VDAC2* in 10q22 at FRA10D (Table 1). Both gene family members include several pseudogenes that are spread all over the genome at other FSs (receiving/accepting FSs).

The birth of new genes is an important feature of genome evolution and an ongoing process. One way to evolve new genes is duplication, with subsequent development of sequence divergence and accumulation of mutations [76, 77]. An alternative mechanism is the de novo formation of new genes, having typically no homologous sequences in the genome [78]. These "orphan" genes are simply structured, small in size, expressed in one tissue [79] and appear as a result of stochastic transcription events in the genome [80]. Overall, CNV formation seems to be a common link enabling fast evolutionary

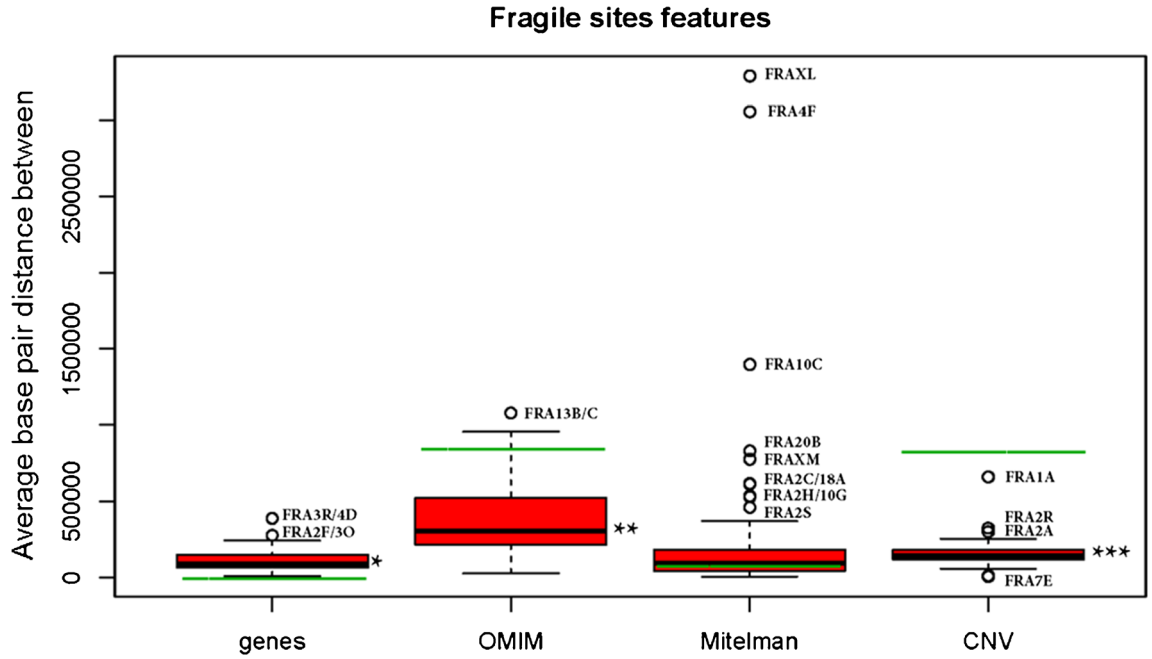

Fig. 1 Sequence features of published fine mapped FSs compared to genome average (green line, GRCh37/hg19) concerning the average base pair distance of genes and transcripts in disease causing OMIM genes, tumor-associated breakpoints from Mitelman database, and CNVs in respect to the FS size

adaption by gene duplication and evolution of new genes within FSs.

These observations led to model of "FS-driven gene and genome evolution," where FSs seem to act as donor and acceptor sites, facilitating the genome wide spreading of gene copies via DNA repair-mediated CNV formation. This sheds new light on the existence and evolutionary conservation of FSs beyond different phylogenetic branches, although their expression can be harmful with respect to aging and cancer on an individual level.

"FS-Driven Gene and Genome Evolution" and Future Directions

Based on this model, several predictions can be made which await further investigations.

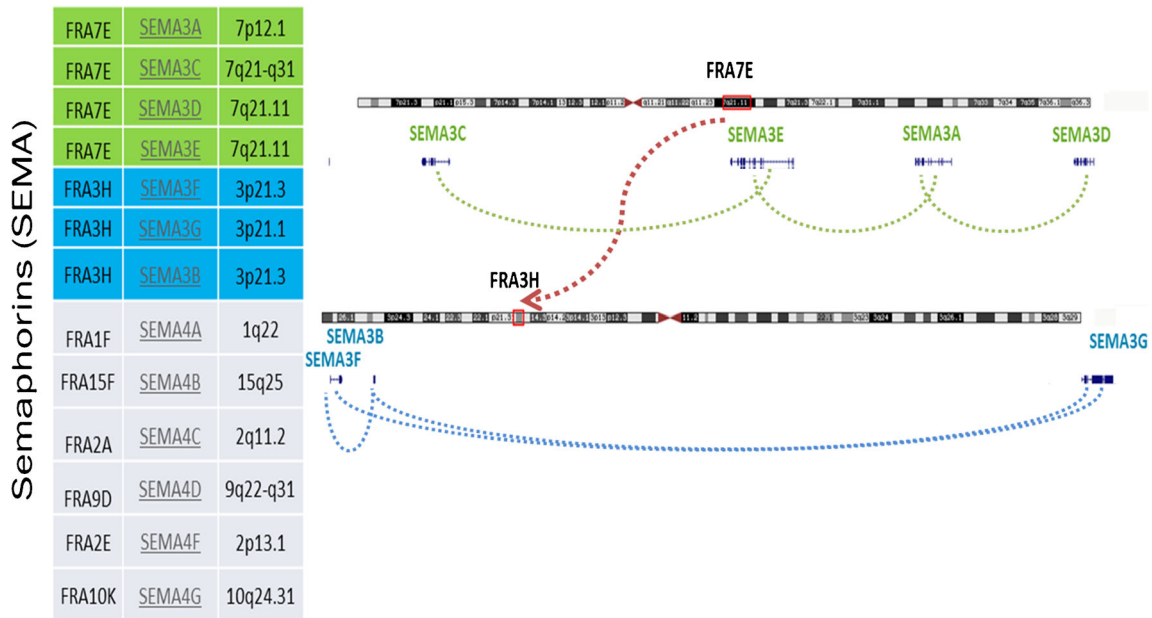

Fig. 2 Genomic distribution of semaphorin gene family members. Four members are located in *cis* in FRA7A (green), three members at FRA3H (blue), and six other members on additional FSs over the genome as indicated

Table 1 Genomic distribution of pseudogene copies from the *VDAC1* and *VDAC2* genes located at different FSs. ID, pseudogene number; Chr, chromosome; FS, fragile site

Parent gene	ID	Chr	Start	End	FS
VDAC1	ENST00000451853.1	13	34,656,566	34,657,447	FRA13A
	ENST00000458323.1	9	97,049,987	97,050,830	FRA9D
	ENST00000423609.1	X	49,397,103	49,397,952	FRAXG
	ENST00000552982.1	12	55,196,530	55,197,373	FRA12A
	ENST00000416715.1	1	215,549,827	215,550,717	FRA1H
	ENST00000447826.1	2	42,690,279	42,691,137	FRA2O
	ENST00000450197.1	1	180,403,935	180,405,071	FRA1G
	ENST00000412766.1	1	157,693,970	157,694,810	FRA1P
VDAC2	ENST00000452925.1	2	135,554,739	135,555,449	FRA2F
	ENST00000467568.1	2	65,432,242	65,433,121	FRA2Q
	ENST00000430995.1	1	118,183,434	118,184,261	FRA1N
	ENST00000462417.1	3	77,365,903	77,366,749	FRA3R
	ENST00000399358.2	21	17,466,735	17,467,692	FRA21A

1. Genomic CNV enrichment sites can predict FSs in silico and can be used for a predictive fine mapping strategy, e.g., via locus-specific FISH probes at the edges of CNV clusters.
2. FSs act as evolutionary flexible and fast-changing regions. This can be checked, e.g., for recent marks of positive selection in population databases, or compared to extinct human species like Neanderthal or Denisova and to recent primate sequence releases [81].
3. FSs from different genomic regions need to be localized in close proximity of DNA repair sites, to facilitate "nonperfect repair," resulting in copies at other, nonhomologs or the same FS in the genome. This could be explored by analyzing sequences around gene family members at different FSs, as those regions were also copied and should have sequence similarities to the parental copy. The divergence might be larger, as there is different selective pressure for coding and noncoding regions as well as for *cis* and *trans* copies [82]. For functionless copies resulting in pseudogenes, one might expect similar degree of divergence for the surrounding sequences.
4. This kind of suggested DNA repair (3.) needs to be located at specific sites in the nucleus and requires DNA mobility. In fact, there is growing evidence on relocation of DNA to specific sites for repair processes to repair centers [83–86]. In support of this, fine mapped FSs can be tracked by FISH in the nucleus of replication stressed cells or chromatin immunoprecipitation (ChIP) and high chromosome conformation capture (Hi-C) analysis of such cells can investigate direct contacts between FSs to prove their proximity during DNA repair processes.

Conclusion

Chromosome breakage in evolution, in clinical cases, in diseases like FA as well to a certain extent in cancer, seems to be a nonrandom process. The common basis of all these breakage events seems to be slightly instable DNA regions, appearing as FSs. As a consequence, this leads to clustering and reuse of breakpoints within the range of FS borders. These predisposed sites might enable genome reshuffling and spreading of gene copies by CNV prone DNA repair, resulting in genome reorganization and variability. On top of that, natural selection can act in evolutionary short terms. However, it is a trade-off and harbors the risk for genomic diseases and cancer. In this context, FSs seem to be the common, yet missing link between breakpoints in disease, aging, cancer and gene/genome evolution being combined in the "FS-driven gene and genome evolution" model. The characterization of FSs is still challenging, but the required basis for comparative studies between the aforementioned groups of break events. Future studies on this topic and the underlying molecular mechanism and consequences will help to advance our understanding of genome dynamics, genome biology, and chromosomal evolution driven by FSs.

Funding Information This work was supported by grants from the IZKF Jena (Interdisciplinary Centre for Clinical Research of Jena University Hospital).

Compliance with Ethical Standards

Conflict of Interest All authors declare no conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Weise A, Kosyakova N, Voigt M, Aust N, Mrasek K, Löhmer S, et al. Comprehensive analyses of white-handed gibbon chromosomes enables access to 92 evolutionary conserved breakpoints compared to the human genome. Cytogenet Genome Res. 2015;45:42–9.
2. Xiaobo F, Pinthong K, Mkrtchyan H, Siripiyasing P, Kosyakova N, Supiwong W, et al. First detailed reconstruction of the karyotype of Trachypithecus cristatus (Mammalia: Cercopithecidae). Mol Cytogenet. 2013;6:58.
3. Fan X, Supiwong W, Weise A, Mrasek K, Kosyakova N, Tanomtong A, et al. Comprehensive characterization of evolutionary conserved breakpoints in four New World Monkey karyotypes compared to Chlorocebus aethiops and Homo sapiens. Heliyon. 2015;1:e00042.
4. Sangpakdee W, Tanomtong A, Fan X, Pinthong K, Weise A, Liehr T. Application of multicolor banding combined with heterochromatic and locus-specific probes identify evolutionary conserved breakpoints in Hylobates pileatus. Mol Cytogenet. 2016;9:17.
5. Schmidt S, Claussen U, Liehr T, Weise A. Evolution versus constitution: differences in chromosomal inversion. Hum Genet. 2005;117:213–9.
6. Weise A, Mrasek K, Klein E, Mulatinho M, Llerena JC Jr, Hardekopf D, et al. Microdeletion and microduplication syndromes. J Histochem Cytochem. 2012;60:346–58.
7. Mulatinho MV, de Carvalho Serao CL, Scalco F, Hardekopf D, Pekova S, Mrasek K, Liehr T, Weise A, Rao N, Llerena JC Jr. Severe intellectual disability, omphalocele, hypospadia and high blood pressure associated to a deletion at 2q22.1q22.3: case report. Mol Cytogenet 2012:5:30.
8. Liehr T, Kosayakova N, Schröder J, Ziegler M, Kreskowski K, Pohle B, et al. Evidence for correlation of fragile sites and chromosomal breakpoints in carriers of constitutional balanced chromosomal rearrangements. Balkan J Med Genet. 2011;14:13–6.
9. Tkach IR, Huleyuk NL, Zastavna DV, Weise A, Liehr T, Ciszkowicz E, et al. Chromosomal aberrations in spontaneously aborted products of conception from Ukraine. Biopolymers Cell. 2017;33:424–33.
10. Li Z, Zhang Q, Mao JH, Weise A, Mrasek K, Fan X, et al. An HDAC1-binding domain within FATS bridges p21 turnover to radiation-induced tumorigenesis. Oncogene. 2010;29:2659–71.
11. Ma K, Qiu L, Mrasek K, Zhang J, Liehr T, Quintana LG, et al. Common fragile sites: genomic hotspots of DNA damage and carcinogenesis. Int J Mol Sci. 2012;13:11974–99.
12. Attie O, Darling AE, Yancopoulos S. The rise and fall of breakpoint reuse depending on genome resolution. BMC Bioinformatics. 2011;12(Suppl 9):S1.
13. Aguado C, Gayà-Vidal M, Villatoro S, Oliva M, Izquierdo D, Giner-Delgado C, et al. Validation and genotyping of multiple human polymorphic inversions mediated by inverted repeats reveals a high degree of recurrence. PLoS Genet. 2014;10:e1004208.
14. • Feng W, Chakraborty A. Fragility extraordinaire: Unsolved mysteries of chromosome fragile sites. Adv Exp Med Biol. 2017;1042:489–526 **Very recent review on still remaining mysteries in FS and their correlation with R-loops and human diseases and includes also the discussion of disease associated rare FS.**
15. Capozzi O, Stanyon R, Archidiacono N, Ishida T, Romanenko SA, Rocchi M. Rapid emergence of independent "chromosomal lineages" in silvered-leaf monkey triggered by Y/autosome translocation. Sci Rep 2018:19;8:3250.
16. • Catacchio CR, Maggiolini FAM, D'Addabbo P, Bitonto M, Capozzi O, Lepore Signorile M, et al. Inversion variants in human and primate genomes. Genome Res. 2018;28:910–20 **Reanalysis of divergent inversions in primate genomes based on most recent genome data at high resolution identified new inversions and confirmed the co localization with recurrent human disease rearrangements.**
17. Le Tallec B, Koundrioukoff S, Wilhelm T, Letessier A, Brison O, Debatisse M. Updating the mechanisms of common fragile site instability: how to reconcile the different views? Cell Mol Life Sci. 2014;71:4489–94.
18. Durkin SG, Glover TW. Chromosome fragile sites. Annu Rev Genet. 2007;41:169–92.
19. Lukusa T, Fryns JP. Human chromosome fragility. Biochim Biophys Acta. 2008;1779:3–16.
20. • Mrasek K, Schoder C, Teichmann AC, Behr K, Franze B, Wilhelm K, et al. Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones. Int J Oncol. 2010;36:929–40 **Catalogues all so far described and 61 new observed FS at cytogenetic level and demonstrated individual differences of FS frequencies.**
21. Le Tallec B, Dutrillaux B, Lachages AM, Millot GA, Brison O, Debatisse M. Molecular profiling of common fragile sites in human fibroblasts. Nat Struct Mol Biol. 2011;18:1421–3.
22. Sarni D, Kerem B. The complex nature of fragile site plasticity and its importance in cancer. Curr Opin Cell Biol. 2016;40:131–6.
23. Fragkos M, Naim V. Rescue from replication stress during mitosis. Cell Cycle. 2017;16:613–33.
24. Boveri T. Zur Frage der Entstehung maligner Tumoren. Jena: Verlag von Gustav Fischer, 1914.
25. • Nalepa G, Clapp DW. Fanconi anaemia and cancer: an intricate relationship. Nat Rev Cancer. 2018;18:168–85 **Reviews the current knowledge on Fanconi anemia and the involvement of 21 causal genes in DNA repair, cancer and FS.**
26. •• Glover TW, Wilson TE, Arlt MF. Fragile sites in cancer: more than meets the eye. Nat Rev Cancer 2017:25;17:489–501. **Recent summary of the molecular features on FS discussed in the focus of cancer.**
27. Thys RG, Lehman CE, Pierce LC, Wang YH. DNA secondary structure at chromosomal fragile sites in human disease. Curr Genomics. 2015;16:60–70.
28. Franchitto A, Pichierri P. Understanding the molecular basis of common fragile sites instability: role of the proteins involved in the recovery of stalled replication forks. Cell Cycle. 2011;10:4039–46.
29. Ozeri-Galai E, Lebofsky R, Rahat A, Bester AC, Bensimon A, Kerem B. Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. Mol Cell. 2011;43:122–31.
30. Shigechi T, Tomida J, Sato K, Kobayashi M, Eykelenboom JK, Pessina F, et al. ATR-ATRIP kinase complex triggers activation of the Fanconi anemia DNA repair pathway. Cancer Res. 2012;72:1149–56.
31. Budzowska M, Kanaar R. Mechanisms of dealing with DNA damage-induced replication problems. Cell Biochem Biophys. 2009;53:17–31.
32. Climprich KA, Cortez D. ATR: an essential regulator of genome integrity. Nat Rev Mol Cell Biol. 2008;9:616–27.
33. Fungtammasan A, Walsh E, Chiaromonte F, Eckert KA, Makova KD. A genome-wide analysis of common fragile sites: what features determine chromosomal instability in the human genome? Genome Res. 2012;22:993–1005.
34. Tsantoulis PK, Kotsinas A, Sfikakis PP, Evangelou K, Sideridou M, Levy B, et al. Oncogene-induced replication stress preferentially

targets common fragile sites in preneoplastic lesions. A genome-wide study. Oncogene. 2008;27:3256–64.
35. Kaiser VB, Semple CA. Chromatin loop anchors are associated with genome instability in cancer and recombination hotspots in the germline. Genome Biol 2018:30;19:101.
36. Palovcak A, Liu W, Yuan F, Zhang Y. Maintenance of genome stability by Fanconi anemia proteins. Cell Biosci. 2017;7:8.
37. Schwab RA, Nieminuszczy J, Shah F, Langton J, Lopez Martinez D, Liang CC, et al. The Fanconi Anemia pathway maintains genome stability by coordinating replication and transcription. Mol Cell. 2015;60:351–61.
38. Ginno PA, Lott PL, Christensen HC, Korf I, Chédin F. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. Mol Cell. 2012;45:814–25.
39. Aguilera A, García-Muse T. R loops: from transcription byproducts to threats to genome stability. Mol Cell. 2012;46:115–24.
40. Costantino L, Koshland D. The Yin and Yang of R-loop biology. Curr Opin Cell Biol. 2015;34:39–45.
41.•• Wilson TE, Arlt MF, Park SH, Rajendran S, Paulsen M, Ljungman M, et al. Large transcription units unify copy number variants and common fragile sites arising under replication stress. Genome Res. 2015;25:189–200 **High resolution induction of CNVs at FS under replication stress demonstrating also a high locus and tissue specificity**.
42. Savelyeva L, Brueckner LM. Molecular characterization of common fragile sites as a strategy to discover cancer susceptibility genes. Cell Mol Life Sci. 2014;71:4561–75.
43. Arlt MF, Wilson TE, Glover TW. Replication stress and mechanisms of CNV formation. Curr Opin Genet Dev. 2012;22:204–10.
44. Song X, Beck CR, Du R, Campbell IM, Coban-Akdemir Z, Gu S, et al. Predicting human genes susceptible to genomic instability associated with Alu/Alu-mediated rearrangements. Genome Res. 2018;28:1228–42.
45. Liu P, Yuan B, Carvalho CMB, Wuster A, Walter K, Zhang L, et al. An organismal CNV mutator phenotype restricted to early human development. Cell. 2017;168:830–42.e7.
46. Magaard Koldby K, Nygaard M, Christensen K, Christiansen L. Somatically acquired structural genetic differences: a longitudinal study of elderly Danish twins. Eur J Hum Genet. 2016;24:1506–10.
47. Mkrtchyan H, Gross M, Hinreiner S, Polytiko A, Manvelyan M, Mrasek K, et al. Early embryonic chromosome instability results in stable mosaic pattern in human tissues. PLoS One. 2010;5:e9591. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0009591.
48. Mkrtchyan H, GrossM HS, Polytiko A, Manvelyan M, Mrasek K, Kosyakova N, et al. The human genome puzzle - the role of copy number variation in somatic mosaicism. Current Genomics. 2010;11:426–31.
49. Bose P, Hermetz KE, Conneely KN, Rudd MK. Tandem repeats and G-rich sequences are enriched at human CNV breakpoints. PLoS One. 2014;9:e101607.
50. Jahic A, Erichsen AK, Deufel T, Tallaksen CM, Beetz C. A polymorphic Alu insertion that mediates distinct disease-associated deletions. Eur J Hum Genet. 2016;24:1371–4.
51. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, et al. Mosaic copy number variation in human neurons. Science. 2013;342:632–7.
52. Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell. 1997;89:215–25.
53. Tanaka SS, Mitsuda SH, Shimizu N. How a replication origin and matrix attachment region accelerate gene amplification under replication stress in mammalian cells. PLoS One. 2014;9:e103439.
54. Yun J, Song SH, Kang JY, Park J, Kim HP, Han SW, et al. Reduced cohesin destabilizes high-level gene amplification by disrupting pre-replication complex bindings in human cancers with chromosomal instability. Nucleic Acids Res. 2016;44:558–72.
55. Le Tallec B, Millot GA, Blin ME, Brison O, Dutrillaux B, Debatisse M. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. Cell Rep. 2013;4:420–8.
56. Hazan I, Hofmann TG, Aqeilan RI. Tumor suppressor genes within common fragile sites are active players in the DNA damage response. PLoS Genet. 2016;12:e1006436.
57. Kotsantis P, Petermann E, Boulton SJ. Mechanisms of oncogene-induced replication stress: Jigsaw falling into place. Cancer Discov. 2018;8:537–55.
58. Macheret M, Halazonetis TD. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. Nature. 2018;555:112–6.
59. Miron K, Golan-Lev T, Dvir R, Ben-David E, Kerem B. Oncogenes create a unique landscape of fragile sites. Nat Commun. 2015;6:7094.
60. Ullastres A, Farré M, Capilla L, Ruiz-Herrera A. Unraveling the effect of genomic structural changes in the rhesus macaque—implications for the adaptive role of inversions. BMC Genomics. 2014;15:530.
61. Pevzner P, Tesler G. Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. Proc Natl Acad Sci U S A. 2003;100:7672–7.
62. Froenicke L. Origins of primate chromosomes—as delineated by Zoo-FISH and alignments of human and mouse draft genome sequences. Cytogenet Genome Res. 2005;108:122–38.
63. Ruiz-Herrera A, Castresana J, Robinson TJ. Is mammalian chromosomal evolution driven by regions of genome fragility? Genome Biol. 2006;7:R115.
64. Farré M, Robinson TJ, Ruiz-Herrera A. An Integrative Breakage Model of genome architecture, reshuffling and evolution: The Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity. Bioessays. 2015;37:479–88.
65. Helmrich A, Stout-Weider K, Hermann K, Schrock E, Heiden T. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. Genome Res. 2006;16:1222–30.
66. Helmrich A, Stout-Weider K, Matthaei A, Hermann K, Heiden T, Schrock E. Identification of the human/mouse syntenic common fragile site FRA7K/Fra12C1—relation of FRA7K and other human common fragile sites on chromosome 7 to evolutionary breakpoints. Int J Cancer. 2007;120:48–54.
67.• Pentzold C, Shah SA, Hansen NR, Le Tallec B, Seguin-Orlando A, Debatisse M, et al. FANCD2 binding identifies conserved fragile sites at large transcribed genes in avian cells. Nucleic Acids Res. 2018;46:1280–94 **By FANCD2 ChIP-seq these authors mapped FS in avian cells and found hinds for a biological conserved function.**
68. Dennis MY, Eichler EE. Human adaptation and evolution by segmental duplication. Curr Opin Genet Dev. 2016;41:44–52.
69. Gazave E, Darré F, Morcillo-Suarez C, Petit-Marty N, Carreño A, Marigorta UM, et al. Copy number variation analysis in the great apes reveals species-specific patterns of structural variation. Genome Res. 2011;21:1626–39.
70. Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW, et al. A burst of segmental duplications in the genome of the African great ape ancestor. Nature. 2009;457:877–81.
71. Song X, Beck CR, Du R, Campbell IM, Coban-Akdemir Z, Gu S, et al. Predicting human genes susceptible to genomic instability associated with Alu/Alu-mediated rearrangements. Genome Res. 2018;28:1228–42.
72. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013;45:1134–40.

73. Marczok S, Bortz B, Wang C, Pospisil H. Comprehensive analysis of genome rearrangements in eight human malignant tumor tissues. PLoS One. 2016;11:e0158995.
74. Wei PC, Lee CS, Du Z, Schwer B, Zhang Y, Kao J, et al. Three classes of recurrent DNA break clusters in brain progenitors identified by 3D proximity-based break joining assay. Proc Natl Acad Sci U S A. 2018;115:1919–24.
75. Wei PC, Chang AN, Kao J, Du Z, Meyers RM, Alt FW, et al. Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. Cell. 2016;164:644–55.
76. Haldane JBS. The part played by recurrent mutation in evolution. Am Nat. 1933;67:5–9.
77. Ohno S. Evolution by gene duplication. Berlin: Springer Verlag; 1970.
78. Ohno S. Birth of a unique enzyme from an alternative reading frame of the preexisted, internally repetitious coding sequence. Proc Natl Acad Sci U S A. 1984;81:2421–5.
79. Schlötterer C. Genes from scratch—the evolutionary fate of de novo genes. Trends Genet. 2015;31:215–9.
80. Ruiz-Orera J, Hernandez-Rodriguez J, Chiva C, Sabidó E, Kondova I, Bontrop R, et al. Origins of de novo genes in human and chimpanzee. PLoS Genet. 2015;11:e1005721.
81. Kronenberg ZN, Fiddes IT, Gordon D, Murali S, Cantsilieris S, Meyerson OS, et al. High-resolution comparative analysis of great ape genomes. Science. 2018;360(6393):eaar6343.
82. Lan X, Pritchard JK. Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. Science. 2016;352:1009–13.
83. Schrank BR, Aparicio T, Li Y, Chang W, Chait BT, Gundersen GG, et al. Nuclear ARP2/3 drives DNA break clustering for homology-directed repair. Nature. 2018;559:61–6.
84. Haber JE. DNA repair: the search for homology. Bioessays. 2018;40:e1700229.
85. Amaral N, Ryu T, Li X, Chiolo I. Nuclear dynamics of heterochromatin repair. Trends Genet. 2017;33:86–100.
86. Strecker J, Gupta GD, Zhang W, Bashkurov M, Landry MC, Pelletier L, et al. DNA damage signalling targets the kinetochore to promote chromatin mobility. Nat Cell Biol. 2016;18:281–90.