Predominance of constitutional chromosomal rearrangements in human chromosomal fragile sites

Inês J. Sequeira[1], João T. Mexia[1], João Santiago[2], Rita Mamede[2], Elisa Silva[2], Jorge Santos[2], Daniel Faria[2], José Rueff[2], Aldina Brás[2*]

[1]Department of Mathematics and CMA, Faculty of Sciences and Technology, Universidade Nova de Lisboa, Caparica, Portugal
[2]Department of Genetics and CIGMH, Faculty of Medical Sciences, Universidade Nova de Lisboa, Lisbon, Portugal
Email: *aldina.gene@fcm.unl.pt

ABSTRACT

Chromosomal fragile sites (CFSs) are loci or regions susceptible to spontaneous or induced occurrence of gaps, breaks and rearrangements. In this work, we studied the data of 4535 patients stored at DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources). We mapped fragile sites to chromosomal bands and divided the 23 chromosomes into fragile and non-fragile sites. The frequency of rearrangements at the chromosomal location of clones found to be deleted or duplicated in the array/CGH analysis, provided by DECIPHER, was compared in Chromosomal Fragile Sites vs. non-Fragile Sites of the human genome. The POSSUM Web was used to complement this study. The results indicated 1) a predominance of rearrangements in CFSs, 2) the absence of statistically significant difference between the frequency of rearrangements in common CFSs vs. rare CFSs, 3) a predominance of deletions over duplications in CFSs. These results on constitutional chromosomal rearrangements are evocative of the findings previously reported by others relatively to cancer supporting the current line of evidence and suggesting that a common mechanism can underlie the generation of constitutional and somatic rearrangements. The combination of insights obtained from our results and their interrelationships can indicate strategies by which the mechanisms can be targeted with preventive medical interventions.

---

*Corresponding author.

1. INTRODUCTION

From a mere cytogenetic observation of 40 years ago [1], fragile sites of human chromosomes became the subject of important studies in human genetics.

Based on their inheritance patterns and population frequency, chromosomal fragile sites (CFSs) were classified in two main categories: rare and common. A subsequent subdivision was made based on the class of chemicals which may induce these types of CFSs [2]. Recently, aphidicolin was able to induce all types of rare and common CFSs, suggesting that these breakage-prone regions are less dependent on specific inducing chemicals than originally considered [3].

Despite the DNA sequence is substantially different among the types of fragile sites, the idea that a general mechanism of failure of replication and fragility at different types of chromosomal fragile sites is emerging [4]. Together the different DNA sequences in fragile sites, including the CCG/CGG trinucleotide repeats, AT-rich minisatellite repeats or AT-dinucleotide-rich islands, are 1) prone to form stable DNA secondary structures which may interfere with DNA replication, 2) have been shown to contain highly flexible DNA sequences that could prevent the replication fork progression and affect chromatin organization, 3) were found to disfavour nucleosome assembly [4].

The traditional classification of fragile sites has recently been questioned [3] and the extension of the definition of CFSs to chromosomal fragile regions has been proposed, since molecular genetic mapping data indicated that breaks do not occur in a defined sequence of CFSs, but most likely in regions prone to breakage as large as 10 Mb [5,6].

Although several lines of evidence indicated that somatic rearrangements occurring within CFSs are associated with cancer development [7], fragile sites have

rarely drawn attention as genomic structures associated with constitutional chromosomal rearrangements leading to birth defects. Chromosome breakage at or near the rare fragile site FRA11B has been implicated in Jacobsen syndrome and an aphidicolin-inducible fragile site, FRA 18C, was identified in the father of a patient with an 18q22.2-qter truncation and the Beckwith-Wiedemann syndrome [8,9].

Given the paucity of data indicating a role for other fragile sites in the formation of constitutional rearrangements as well as a study reporting on the lack of association between fragile sites and constitutional chromosome breakpoints [10], it was thought that the involvement of CFSs might be minimal in constitutional aberrations. The recently obtained evidence that chromosomal instability associated with CFSs plays an important role in gross deletions and duplications in germ cell lines, considered causal in human diseases [11], led us readdress this old issue.

Therefore, the main aim of our study was to determine the frequency of constitutional chromosomal rearrangements in CFSs, including common and rare CFSs. Comparison of the frequencies of chromosomal aberrations in these sites will help clarify the molecular events that might lead to disease.

2. METHODS

2.1. Data

The genomic position of fragile and non-fragile sites was based on the NCBI Map Viewer (Build 37.2). The consented DECIPHER patient data, including the start and the end positions of the deleted or duplicated regions, remapped to GRCh37 (hg19), were provided by DECIPHER in 2011. The analysis of a total of 4535 DECIPHER patients was performed. The POSSUM Web was assessed during 2010-2011. Only rearrangements involving DNA fragments larger than 1 Kb were considered.

2.2. Study Strategy

The human genome was divided into regions by inspecting their sequential positioning, in line with Laganà *et al.* [12]. Two sequential bands associated with fragile sites are grouped together to form a FR and the region between two separate FRs is considered a non-Fragile Region. As such 258 different regions (fragile and non-fragile) could be considered and were used throughout this study. We take into account that 1) the fragile sites' set was recently extended [3], 2) common and rare chromosomal fragile sites share some characteristics [4], 3) some rare fragile sites span the same genomic regions as common fragile sites [13], and 4) the CFS do not break at defined sequences but in breakage-prone regions [5,6]. The Y chromosome was excluded from this study, since

there is only one suggestion that this chromosome might contain a fragile site [14].

2.3. Statistical Analysis

For each chromosome, the number of deleted and duplicated regions was obtained, including the start and end positions, as well as the lengths, of fragile l_{FR} and non-fragile $l_{non\text{-}FR}$ regions. Rearrangements could be grouped into four classes: N_1 (Yes-Yes) as the proportion of rearrangements that start and end in FRs, N_2 (Yes-Start) as the proportion of rearrangements that start in FRs and end in non-FRs, N_3 (Yes-End) as the proportion of rearrangements that start in non-FRs and end in FRs and N_4 (No) as the proportion of rearrangements that start and end in non-FRs.

We used these values to calculate the rearrangement intensities given by $i_1 = N_1/l_{FR}$, $i_2 = N_2/\left(l_{FR} \times l_{non\text{-}FR}\right)^{1/2}$ $i_3 = N_3/\left(l_{FR} \times l_{non\text{-}FR}\right)^{1/2}$ and $i_4 = N_4/l_{non\text{-}FR}$. As such, the intensity was defined as the frequency of rearrangements occurring in each class, weighted by the respective length of the region. Intensity values were plotted, locating the different chromosomes for each pair of intensities.

The graphs correspond to the pairs of variables (x,y) where each chromosome will be represented by a point. We plotted the straight line $y = x$ to see if we had more chromosomes with $y > x$ or more chromosomes with $y < x$. In order to compare two different intensities the sign test was used.

Moreover, to position the chromosomes when considering the four intensities, we carried out a principal component analysis (PCA). The leading principal components are the linear combinations of the initial variables containing more information. We used the first and second principal components to get a global representation of rearrangement occurrence in all chromosomes.

3. RESULTS AND DISCUSSION

3.1. Predominance of Constitutional Chromosomal Rearrangements in CFSs

Figure 1 indicated that chromosomal rearrangements, both deletions and duplications, responsible for chromosomal imbalance and associated phenotype alterations, occur predominately in CFSs. This predominance of rearrangements in CFSs was also found in the POSSUM syndromes database (data not shown). This conclusion is in line with the current evidence, and definitely away from the finding that there was no particular association between fragile sites and constitutional chromosome rearrangements [10]. The clarification of this point may be due, at least in part, to advances in cytogenetic techniques from the conventional G banding to the recent
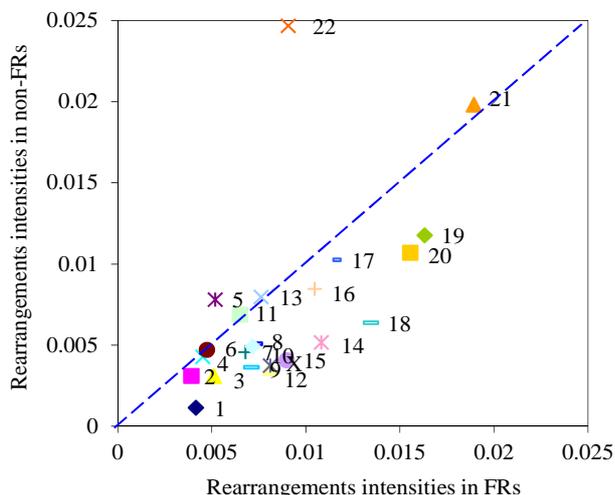
Figure 1. Representation of the intensity of rearrangement located in FRs vs. intensity of rearrangement located in non-FRs. Each chromosome is represented by a point. The coordinates of each point are a pair of intensities of rearrangement located in FRs and located in non-FRs. The straight line ($y = x$) allow us to identify more chromosomes with $x > y$. To test the hypothesis of equal intensities against the alternative of higher intensities for fragile regions we used the sign test, given its robustness. We obtained a *p-value* = 0.005 so we can clearly reject the hypothesis of equal intensities at the 1% level.

array-based techniques.

When the CGH array has been introduced in the clinical practice, it became increasingly clear that the diagnostic potential of this technology was greater than G banding [15,16]. Reviewing 29 studies of patients with developmental delay/mental retardation (DD/MR), Hochstenbach *et al.* [16] showed that a yield of approximately twice the rate of the classical cytogenetics' findings would be achieved using array analysis.

On the other hand, our results indicating predominance of constitutional chromosomal rearrangements in CFSs are evocative of the findings previously reported by others concerning rearrangements in cancer. Namely, Burrow *et al.* [17] reported that most of the breakpoints in pairs of genes involved in cancer-specific recurrent translocations are located in human chromosomal fragile sites, supporting a causal role for fragile sites in the generation of chromosomal rearrangements in somatic cells. Using a custom-designed high-density CGH analysis to study the junction sequences of approximately 500 breakpoints in germ cell lines and cancer cell lines involving PARK2 or DMD, Mitsui *et al.* [11,18] suggested that a common mechanism may be involved in the generation of rearrangements in both types of cell lines. Our results extend the findings of these authors, adding evidence that chromosomal fragility associated with CFSs plays a role in constitutional chromosomal rearrangements as well.

In what concerns the exception of chromosome 22

(**Figures 1** and **2**), it was reported that the 22q11.2 region, a hotspot for chromosomal rearrangements, showed instability features of fragile sites [19], however this region is not yet classified as a CFS and as such not included in FRs in this study.

3.2. Absence of Significant Statistical Difference between Rare and Common CFSs Relative to the Frequency of Chromosomal Rearrangements

In spite of the useful dichotomy between rare and common chromosomal fragile sites, studies showed that, fragile sites have actually a broad continuous gradient of frequency ranging from very rare to very common (for discussion see [20]). Both types of fragile sites display common molecular characteristics associated with chromosomal rearrangements, both *in vitro* and *in vivo* (for review see [4]). Supporting this association *in vitro* is the fact that following induction of the fragile site, a proportion of cells from individuals with rare fragile sites are found to have various duplications or deletions of material distal to the fragile site [21]. This is considered to be the result of breakage at the fragile site followed by nondisjunction of the distal chromosomal material [21]. Also, common fragile sites have been shown to display a number of characteristics of unstable and highly recombinogenic DNA *in vitro*, including chromosome rearrangements [22].

In vivo evidence of instability and constitutional chromosomal breakage is given by the chromosomal deletions in a proportion of patients with Jacobsen and Fragile X syndromes [23-25], as well as the association between common fragile sites and chromosomal deletions and translocations occuring in human genetic disorders [20].
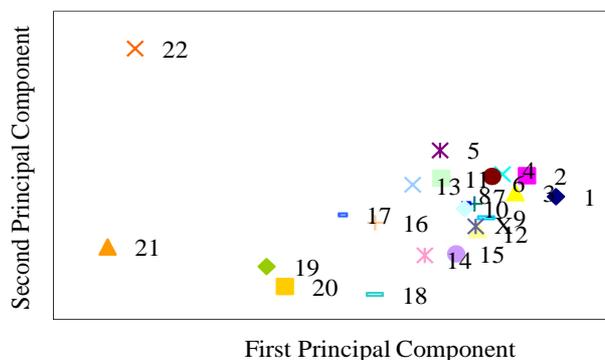

Figure 2. First Principal Component (PC1) vs. Second Principal Component (PC2). The fraction of the total information contained in PC1 is 79% and by PC2 is 18%. Intuitively, the first two PCs show that chromosome 22 is isolated. The isolation of chromosome 22 is also evident when we include the cases in which only the start position or only the end position of deleted or duplicated regions are located in FRs.

Since common and rare fragile sites are breakage-prone regions and recent data suggest that the differences between mechanisms of instability at common versus rare fragile sites are not so stringent [3], we compared the frequency of rearrangements in rare vs. common fragile sites. We could not detect any statistically significant difference between rare and common fragile sites relative to the frequency of chromosomal rearrangements (**Figure 3**). For this lack of statistical significance between rare and common fragile sites, we pursued this study without discriminating between common and rare fragile sites.

3.3. The Most Frequent Chromosomal Rearrangements That Occur in CFSs Are Deletions

A higher number of deletions compared to duplications in CFSs was found (**Figure 4**) but not in non-FRs. These findings can be due to the differences in mechanisms of generation of these chromosomal rearrangements. The NAHR mechanism favours deletions over duplications, because deletions can result from crossovers both in cis and in trans, whereas duplications can only result from crossovers in trans [26]. As the majority (66.7%) of the NAHR-prone regions described by Liu *et al.* (2012) are wholly or partially included in CFSs, it is possible that this mechanism play a role in the rearrangements occurring in CFSs. In the male germline, it has been found that deletions occur approximately twice as frequently as duplications on autosomes [27]. In spite of the role of selection in the population, it is possible that other factors are involved in this higher rate of deletions [27].

Among many factors determining the fragility of CFSs, changes in replication time of DNA seem to play an important role. The breakpoint-clustering region is replicated later and flanked by the high-flexibility peaks and the R/G band boundaries [11]. When replication forks slow, the likelihood that replication is incomplete at the time of entry into division is increased in the region without initiating events. This contributes to explain the high frequency of breaks observed in CFSs [28]. Furthermore, the analysis of nucleotide-sequence content flanking the breakpoints in CFSs demonstrated junctions with microhomologies to be predominant, favouring the involvement of MMEJ at CFSs [11]. Studying two Common Fragile-Site-Associated Loci, PARK2 and DMD, in germ cell and cancer cell lines, these authors also found that deletions were more frequently observed than duplications [11]. Our results are consistent with these findings, showing that the higher frequency of deletions versus duplications generation occurs at the expense of deletions at CFSs.

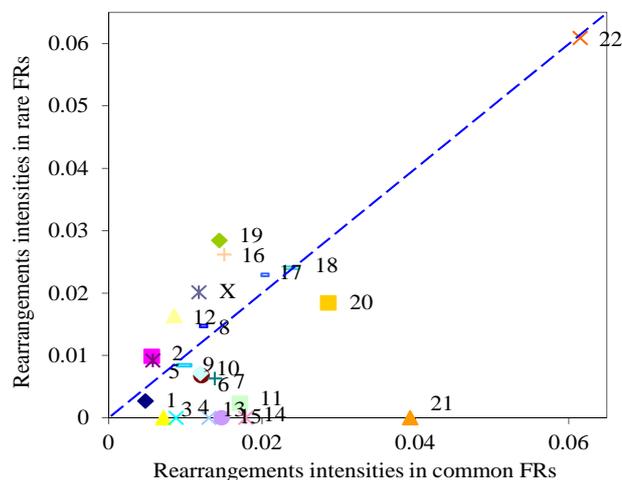Work is in progress to better clarify these mechanisms in chromosomal fragile sites.

Figure 3. Representation of the intensity of rearrangement located in common FRs vs. intensity of rearrangement located in rare FRs in each chromosome. We used again the sign test for comparing the intensities for rare and common fragile regions. The *p-value* obtained was 0.202 so we cannot reject the hypothesis of equal intensities.
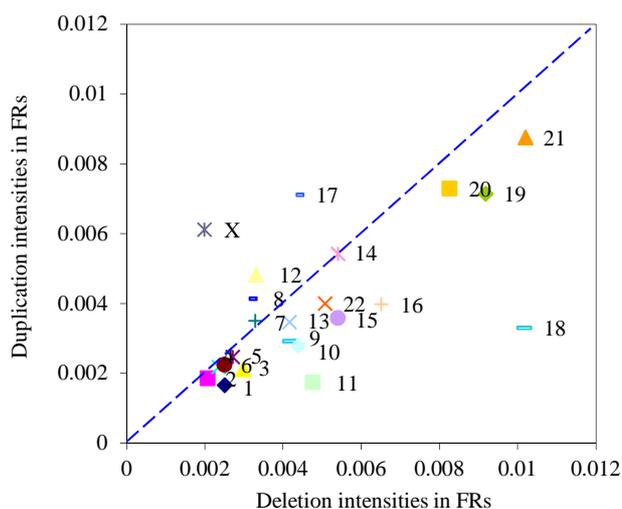

Figure 4. Representation of the intensity of deletions vs. duplications located in FRs in each chromosome. Each chromosome is represented by a point. The coordinates of each point are a pair of intensities of deletions and duplications located in FRs. The straight line $y = x$ allow us to identify more chromosomes with the intensity of deletions located in FRs larger than the intensity of duplications located in FRs. The sign test indicates the rejection, at the 1% level (*p-value* = 0.008), the hypothesis of equal intensities.

4. ACKNOWLEDGEMENTS

REFERENCES

[1] Lubs, H.A. (1969) A marker X chromosome. *American Journal of Human Genetics*, **21**, 231-44.

[2] Speicher, M.R. (2010) Chromosomes. In: Speicher, M. R., Antonarakis, S.E. and Motulsky, A.G., Eds., Vogel and Motulsky's Human Genetics Problems and Approaches, Springer Verlag, Berlin, Heidelberg, 55-138.

[3] Mrasek, K., Schoder, C., Teichmann, A.C., Behr, K., Franze, B., Wilhelm, K., Blaurock, N., Claussen, U., Liehr, T. and Weise A. (2010) Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones. *International Journal of Oncology*, **36**, 929-940.

[4] Lukusa, T. and Fryns, J.P. (2008) Human chromosome fragility. *Biochimica et Biophysica Acta*, **1779**, 3-16. doi:10.1016/j.bbagrm.2007.10.005

[5] Curatolo, A., Limongi, Z.M., Pelliccia, F. and Rocchi, A. (2007) Molecular characterization of the human common fragile site FRA1H. *Genes Chromosomes Cancer*, **46**, 487-493. doi:10.1002/gcc.20432

[6] Zhu, Y., McAvoy, S., Kuhn, R. and Smith, D.I. (2006) RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene*, **25**, 2901-2908. doi:10.1038/sj.onc.1209314

[7] Dillon, L., Burrow, A. and Wang, Y.-H. (2010) DNA instability at chromosomal fragile sites in cancer. *Current Genomics*, **11**, 326-337. doi:10.2174/138920210791616699

[8] Debacker, K. and Kooy, R.F. (2007) Fragile sites and human disease. *Human Molecular Genetics*, **16**, R150-R158.

[9] Debacker, K., Winnepenninckx, B., Ben-Porat, N., Fitz-Patrick, D., Van Luijk, R., Scheers, S., Kerem, B. and Frank Kooy, R. (2007) FRA18C: A new aphidicolin-inducible fragile site on chromosome 18q22, possibly associated with *in vivo* chromosome breakage. *Journal of Medical Genetics*, **44**, 347-352. doi:10.1136/jmg.2006.044628

[10] Mariani, T., Musio, A. and Simi, S. (1995) No statistical association between fragile sites and constitutional chromosome breakpoints. *Cancer Genetics and Cytogenetics*, **85**, 78-81.

[11] Mitsui, J., Takahashi, Y., Goto, J., Tomiyama, H., Ishikawa, S., Yoshino, H., Minami, N., Smith, D.I., Lesage, S., Aburatani, H., Nishino, I., Brice, A., Hattori, N. and Tsuji, S. (2010) Mechanisms of genomic instabilities underlying two common fragile-site-associated loci, PARK2 and DMD, in germ cell and cancer cell lines. *The American Journal of Human Genetics*, **87**, 75-89.

doi:10.1016/j.ajhg.2010.06.006

[12] Laganà, A., Russo, F., Sismeiro, C., Giugno, R., Pulvirenti, A. and Ferro, A. (2010) Variability in the incidence of miRNAs and genes in fragile sites and the role of repeats and CpG islands in the distribution of genetic material. *PLoS One*, **5**, e11166. doi:10.1371/journal.pone.0011166

[13] Zlotorynski, E., Rahat, A., Skaug, J., Ben-Porat, N., Ozeri, E., Hershberg, R., Levi, A., Scherer, S.W., Margalit, H. and Kerem, B. (2003) Molecular basis for expression of common and rare fragile sites. *Molecular and Cellular Biology*, **23**, 7143-7151. doi:10.1128/MCB.23.20.7143-7151.2003

[14] Holden, J., Ridgway, P. and Smith, A. (1986) A possible fragile-site at Yq12: Case report and possible relevance to *de novo* structural rearrangements of the Y-chromosome. *American Journal of Medical Genetics*, **23**, 545-555. doi:10.1002/ajmg.1320230147

[15] Schaaf, C.P., Wiszniewska, J. and Beaudet, A.L. (2011) Copy number and SNP arrays in clinical diagnostics. *Annual Review of Genomics and Human Genetics*, **12**, 25-51. doi:10.1146/annurev-genom-092010-110715

[16] Hochstenbach, R., van Binsbergen, E., Engelen, J., Nieuwint, A., Polstra, A., Poddighe, P., Ruivenkamp, C., Sikkema-Raddatz, B., Smeets, D. and Poot, M. (2009) Array analysis and karyotyping: Workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherland. *European Journal of Medical Genetics*, **52**, 161-169. doi:10.1016/j.ejmg.2009.03.015

[17] Burrow, A.A., Williams, L.E., Pierce, L.C. and Wang, Y.H. (2009) Over half of breakpoints in gene pairs involved in cancer-specific recurrent translocations are mapped to human chromosomal fragile sites. *BMC Genomics*, **10**, 59. doi:10.1186/1471-2164-10-59

[18] Mitsui, J. and Tsuji, S. (2011) Common chromosomal fragile sites: Breakages and rearrangements in somatic and germline cells, atlas of genetics and cytogenetics in oncology and haematology. http://atlasgeneticsoncology.org/Deep/ChromFragSitesID20098.html

[19] Puliti, A., Rizzato, C., Conti, V., Bedini, A., Gimelli, G., Barale, R. and Sbrana, I. (2010) Low-copy repeats on chromosome 22q11.2 show replication timing switches, DNA flexibility peaks and stress inducible asynchrony, sharing instability features with fragile sites. *Mutation Research*, **686**, 74-83. doi:10.1016/j.mrfmmm.2010.01.021

[20] Savelyeva, L., Sagulenko, E., Schmitt, J.G. and Schwab, M. (2006) Low-frequency common fragile sites: Link to neuropsychiatric disorders? *Cancer Letters*, **232**, 58-69. doi:10.1016/j.canlet.2005.08.033

[21] Sutherland, G.R. and Baker, E. (2000) The clinical significance of fragile Sites on human chromosomes. *Clinical Genetics*, **58**, 157-161. doi:10.1034/j.1399-0004.2000.580301.x

[22] Glover, T.W. (1998) Instability at chromosomal fragile sites. *Recent Results in Cancer Research*, **154**, 185-199. doi:10.1007/978-3-642-46870-4_11

[23] Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L., Langdon, W.Y., Sutherland, G. R., Richards, R.I. and Tunnacliffe, A. (1995) Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene CBL2. *Nature*, **376**, 145-149. doi:10.1038/376145a0

[24] Gedeon, A.K., Baker, E., Robinson, H., Partington, M.W., Gross, B., Manca, A., Korn, B., Poustka, A., Yu, S., Sutherland, G.R. and Mulley, J.C. (1992) Fragile X syndrome without CCG amplification has an FMR1 deletion. *Nature Genetics*, **1**, 341-344. doi:10.1038/ng0892-341

[25] Wohrle, D., Kotzot, D., Hirst, M.C., Manca, A., Korn, B., Schmidt, A., Barbi, G., Rott, H.D., Poustka, A., Davies, K.E. and Steibach, P. (1992) A microdeletion of less than 250 kb, including the proximal part of the FMR-I gene and the fragile-X site, in a male with the clinical phenotype of fragile X syndrome. *The American Journal of Human Genetics*, **51**, 299-306.

[26] Liu, P., Carvalho, C.M., Hastings, P.J. and Lupski, J.R. (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Current Opinion in Genetics & Development*, **22**, 211-220. doi:10.1016/j.gde.2012.02.012

[27] Turner, D.J., Miretti, M., Rajan, D., Fiegler, H., Carter, N.P., Blayney, M.L., Beck, S. and Hurles, M.E. (2008) The rates of *de novo* meiotic deletions and duplications causing several genomic disorders in the male germline. *Nature Genetics*, **40**, 90-95. doi:10.1038/ng.2007.40

[28] Letessier, A., Birnbaum, D., Debatisse, M. and Chaffanet, M. (2011) Genome: Does a paucity of initiation events lead to fragility? *Medical Sciences* (*Paris*), **27**, 707-709. doi:10.1051/medsci/2011278011