Human Molecular Genetics, 2000, Vol. 9, No. 19 2879–2884

Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity

Petra G.M. van Overveld[1], Richard J.F.L. Lemmers[1], Giancarlo Deidda[1,2], Lodewijk Sandkuijl[1], George W. Padberg[3], Rune R. Frants[1] and Silvère M. van der Maarel[1,+]

[1]Department of Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, PO Box 9502, 2333 AL Leiden, The Netherlands, [2]National Research Council, Rome, Italy and [3]Department of Neurology, University of Nijmegen, Nijmegen, The Netherlands

Received 4 August 2000; Revised and Accepted 17 September 2000

Chromosomal rearrangements occur more frequently in subtelomeric domains than in other regions of the genome and are often associated with human pathology. To further elucidate the plasticity of subtelomeric domains, we examined the 3.3 kb D4Z4 repeat array on chromosome 4 and its homologue on chromosome 10 in 208 Dutch blood donors by pulsed field gel electrophoresis. These subtelomeric repeats are known to rearrange and partial deletions of this polymorphic array on chromosome 4 are associated with facioscapulohumeral muscular dystrophy (FSHD), an autosomal dominant myopathy. Our results show that mitotic rearrangements occur frequently as 3% of individuals display somatic mosaicism for a repeat expansion or contraction explaining the high variability of subtelomeric repeat array sizes. Translocated 4-type repeat arrays on chromosome 10 and the reverse configuration of 10-type repeat arrays on chromosome 4 are observed in 21% of individuals. The translocated repeat arrays on chromosome 4 tend to be more heterogeneous than 4-type repeats on chromosome 10. The repeat length on chromosome 4 is on average larger than on chromosome 10. But on both chromosomes we observe a multi-modal repeat length distribution with equidistant peaks at intervals of 65 kb, possibly reflecting a higher-order chromatin structure. Interestingly, in as many as six random blood donors (3%) we identified FSHD-sized 4-type repeat arrays. Assuming that these individuals are clinically unaffected, these results imply an incomplete penetrance in the upper range of FSHD alleles. Overall, the observed dynamic characteristics of these homologous domains may serve as a model for subtelomeric plasticity.

INTRODUCTION

The chromosomes of many organisms contain subtelomeric regions, which include single copy sequences, highly repetitive sequences and non-homologous chromosome ends (1,2). These regions lie immediately adjacent to the dynamic $(TTAGGG)_n$ telomere repeat arrays and display a variable size distribution ranging from a few hundred base pairs (e.g. human XpYpter) (3) to >100 kb [4qter (4) and 16pter (5)]. In yeast, chromosome ends are composed of different subtelomeric repeated sequences, such as X and Y′ (6–8) and rearrangements of Y′ elements can easily be detected in mitosis and meiosis (9,10). Data from other systems also indicate that subtelomeric regions of chromosomes are dynamic structures and may in fact support high levels of meiotic recombination in the initial homology searches, which precede chromosome synapsis (11–15).

In 1997, a model was proposed into which subtelomeric regions can be divided in a distal and a proximal subdomain separated by degenerated $(TTAGGG)_n$ repeats (16). The distal subdomain contains repetitive sequences that interact with all chromosome ends, whereas the proximal domain only interacts with a subset of (non-) homologous chromosome ends. It was suggested that subtelomeric regions act as a buffer between functional telomeric repeats and genes proximal of the subtelomeric regions, thereby protecting these genes from heterochromatinization and silencing, as demonstrated in yeast (17–21).

Chromosomal rearrangements occur more frequently in (sub)telomeric domains than in other regions of the genome and often result in disease. Translocations of chromosome ends have been reported to cause for example α-thalassaemia mental retardation syndrome, Wolf–Hirschhorn syndrome and Miller–Dieker syndrome (22–24). Subtelomeric rearrangements may also account for 7% of the idiopathic mental retardation cases and congenital anomalies (25). Probably the most intriguing example of disease-associated subtelomeric rearrangements is seen in facioscapulohumeral muscular dystrophy (FSHD), an autosomal dominant myopathy mapped to 4qter. This proximal subtelomeric domain harbours the

+To whom correspondence should be addressed. Tel: +31 71 527 6107; Fax: +31 71 527 6075; Email: maarel@lumc.nl

polymorphic D4Z4 repeat array, which is composed of 11 up to 150 3.3 kb repeated elements in normal individuals. FSHD patients carry a repeat array of <11 units, due to a partial deletion of the D4Z4 repeat array (26,27).

The subtelomere of chromosome 10 is highly similar to 4q35 and also contains a polymorphic repeat array highly homologous to D4Z4 (28,29). A specific *Bln*I site within each repeat unit of chromosome 10 allows discrimination between both chromosomes (30). In a survey of 50 healthy Dutch male controls we could demonstrate the presence of chromosome 10-type repeats on chromosome 4 in 10% of the Dutch population. The reverse configuration of chromosome 4-type repeats on chromosome 10 is equally frequent (31). Hybrid repeats, consisting of both 4-type and 10-type repeat units, have also been identified. Several studies demonstrated that only short repeat arrays on chromosome 4 are causally related to FSHD (32,33).

Recently, we demonstrated that the partial D4Z4 repeat array deletion in FSHD arises mitotically in at least 40% of cases and that the presence of 4-type repeat arrays on chromosome 10 in these mosaic individuals is increased by almost five times (34). We hypothesized that the FSHD deletion occurs mainly by somatic interchromosomal gene conversion in which the presence of a fully homologous repeat array on a non-homologous chromosome is a predisposing factor.

In this study, we examined the subtelomeric repeat configurations on chromosomes 4 and 10 of 208 unrelated healthy individuals of a Dutch control population. We confirmed the 20% exchanges between 4q35 and 10q26 for both sexes and demonstrated a significant difference in repeat array length between both these chromosomes. Preferred repeat array lengths were observed possibly reflecting a higher order chromatin structure. Our results indicate a high incidence of mitotic rearrangements as part of the dynamic subtelomere characteristics. This study sheds new light on subtelomeric plasticity and may serve as a model for other subtelomeric loci.

RESULTS

Repeat array configurations

We examined the subtelomeric repeat array configurations on chromosomes 4q35 and 10q26 in 208 unrelated healthy individuals: 128 males and 80 females. Alleles were sized by pulsed field gel electrophoresis (PFGE) and assigned to their chromosomal origin based on their *Bln*I sensitivity. The various allele configurations were classified as indicated in Table 1. According to the Mann–Withney *U*-test, no significant differences between repeat array distribution of males and females were observed, neither for chromosome 4 ($P = 0.84$) nor chromosome 10 ($P = 0.67$). Accordingly, data for all individuals were pooled for further analyses.

Most of the individuals (76%) displayed a standard pattern of 4-type arrays on chromosome 4 and 10-type arrays on chromosome 10, designated as class A. The remaining individuals (21%) displayed non-standard configurations with translocated repeat arrays.

Table 1. Subtelomeric repeat array constitutions on chromosomes 4q and 10q of 208 unrelated healthy Dutch individuals

	chr.4	chr.4	chr.10	chr.10	class	n
standard	□▷▷▷▷	□▷▷▷▷	■▶▶▶▶	■▶▶▶▶	A	158
10 on 4	•	□▶▶▶▶	•	•	B	6
	•	□▷▷▷▷	•	•	C	5
	•	□▶▶▷▷	•	•	D	7
	□▶▶▶▶	□▶▶▶▶	•	•	E	0
4 on 10	•	•	■▷▷▷▷	•	F	18
	•	•	■▶▶▷▷	•	G	4
	•	•	■▷▷▶▶	•	H	1
	•	•	■▷▷▷▷	■▷▷▷▷	I	2
complex allele constitution					X	1
somatic mosaicism					M	6

The region hybridizing to p13E-11 is represented by a square whereas 4-type and 10-type repeat arrays are indicated by four white and black triangles, respectively. Dots indicate a similar repeat structure as that at the top of the column. Individuals are classified in 11 different groups.
Class A, the most common allele configuration, is a standard repeat array distribution with 4-type arrays on chromosome 4 and 10-type arrays on chromosome 10.
Classes B–E are individuals that carry a 10-type repeat array on chromosome 4, either homogeneous (B and E) or heterogeneous (C and D). The same holds true for the individuals carrying either homogeneous (F and I) or heterogeneous (G and H) 4-type repeats on chromosome 10.
Mosaic individuals are scored as M, whereas one individual had such complex repeat array structure (X) that we could not assign this individual to one of the previously defined classes. The last column (*n*) indicates the number of individuals in each class.

Repeat size distribution

Only alleles of class A individuals were included in the analysis of allele size distribution (Fig. 1) since this class represents the normal unbiased allele size distribution, whereas in other classes the allele sizes may be influenced by the translocated repeat arrays. All alleles of class A individuals were grouped into intervals of 20 kb each.

Excluding allele sizes of <38 kb (associated with FSHD when residing on chromosome 4), the median of 4-type repeat arrays is 96 kb, whereas the median of 10-type repeat arrays is 75 kb. After correction for the difference in median repeat size, no significant difference was detected between both repeat size distributions according to the Kolmogorov–Smirnov test. As plotted in Figure 1, the repeat sizes on both chromosomes do not display a uniform distribution. The observed distributions for chromosomes 4 and 10 both agreed significantly better with a mixture of three normal distributions with equidistant means than with a mixture of two normal distributions ($\chi^2 = 27.25$, $P < 10^{-6}$ and $\chi^2 = 19.76$, $P < 10^{-5}$, for chromosomes 4 and 10, respectively). No significant further improvement of the likelihoods was obtained when a mixture of three normal distributions with non-equidistant means was fitted ($\chi^2 = 0.54$, $P \approx 0.46$ and $\chi^2 = 0.84$, $P \approx 0.36$ for chromosomes 4 and 10). The most parsimonious model therefore indicates that the allele sizes follow distributions with periodicity.

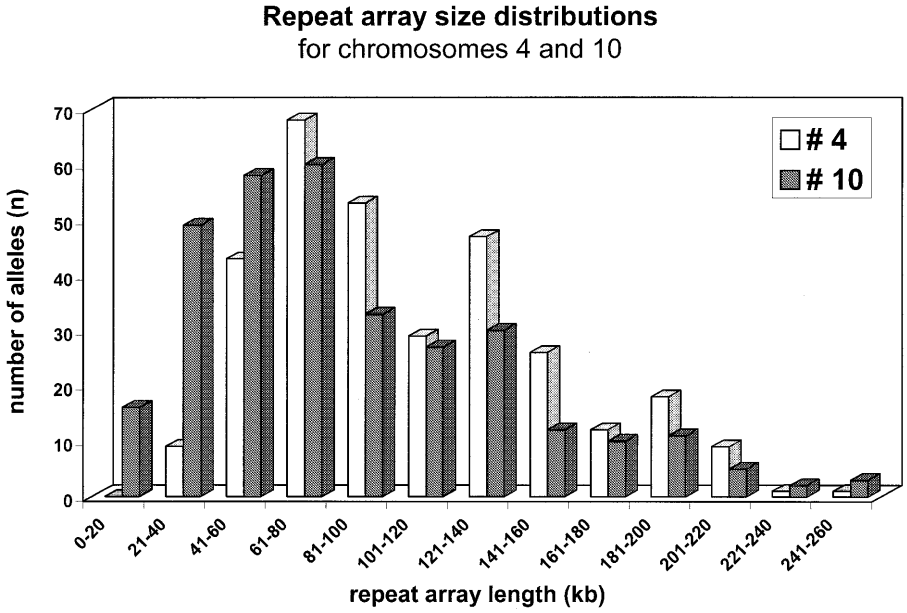

Figure 1. Repeat array size distribution of 4-type and 10-type repeat arrays from individuals in class A. On the *x*-axis the repeat array length is plotted in intervals of 20 kb. The *y*-axis displays the number of alleles of class A in each size interval. Chromosome 4-type repeat arrays are represented by open bars, whereas 10-type repeats are represented by black bars.

FSHD-sized repeat arrays

Repeat arrays <38 kb and residing on chromosome 4 are associated with FSHD. In this survey, we identified 4-type repeat arrays of <38 kb in six individuals (25, 25, 30, 32, 35 and 35 kb). One of the individuals carrying the 30 kb 4-type repeat array was classified as F, carrying only 4-type repeat units on chromosome 10. The other individuals carried a standard repeat array configuration (class A).

Homogeneity of translocated repeat arrays

Twenty-one per cent of individuals displayed a non-standard repeat array configuration on chromosomes 4 and 10. Ten-type repeat arrays on chromosome 4, termed B–E, were identified in 18 individuals (9%), whereas the reverse configuration i.e. 4-type repeat arrays on chromosome 10, termed F–I, was present in 25 individuals (12%) (Table 1). One individual (class X) had such complex repeat array configuration that it could not be assigned to one of the previously defined classes.

Studying all alleles in individuals carrying translocated repeat arrays (36 in classes B–E and 50 in classes F–I), 10-type arrays on chromosome 4 tended to be more heterogeneous than 4-type arrays on chromosome 10. Of the translocated 10-type arrays on chromosome 4 (classes B–E), 17% (6/36) were composed of a homogeneous array of 10-type repeat units whereas 33% (12/36) were hybrid arrays, consisting of clusters of 4-type and 10-type repeat units. Conversely, 44% (22/50) of chromosomes 10 carried a homogeneous translocated array of 4-type units only, whereas 10% (5/50) were hybrid arrays.

There was a significant difference in homogeneity of translocated repeat arrays between classes D and H ($P = 0.006$). These classes carry a comparable repeat array structure: a chromosome 4 backbone with 10-type repeats followed by a cluster of 4-type repeats (class D) and a chromosome 10 backbone with 4-type repeats followed by a cluster of 10-type repeats (class H). The other types of comparable configuration (classes C and G) did not differ significantly.

Somatic mosaicism

In this survey, somatic mosaicism (defined as a fifth repeat array fragment on PFGE) was observed in 3% of individuals (6/208, 4 males and 2 females). Five individuals displayed somatic mosaicism with a standard allele configuration whereas the sixth individual carried an extra 4-type repeat array on chromosome 10. Two of these individuals were mosaic for chromosome 4 and three for chromosome 10-type arrays. The presence of an extra 4-type repeat in the sixth individual did not allow us to assign the arrays to their respective chromosomes.

DISCUSSION

In this survey, we analysed the subtelomeric D4Z4 repeat array configuration of the human chromosomes 4qter and 10qter in 208 individuals of the Dutch population. This study corroborates an earlier study (31) and provides new insights into the behaviour of these subtelomeric domains.

The end of human chromosome 4q displays a similar structural organization as reported for 4pter, 16pter and 22qter (16). The distal subdomain on 4qter contains repetitive sequences as found on many chromosome ends (J. Hewitt, personal communication), whereas the proximal domain is largely comprised of the D4Z4 repeat array studied here.

The proximal and distal subtelomeric domains on chromosomes 4qter and 10qter are >95% homologous. The distal subdomain comprises 25 kb whereas the proximal domain may vary from 50 to >500 kb due to the polymorphic 3.3 kb repeat

array. Despite their high homology, 10-type repeat units can be discriminated from 4-type repeat units by an internal *Bln*I site (30). This *Bln*I-based array discrimination makes these subtelomeres ideal models to study subtelomere plasticity. Employing PFGE, we as well as others have already demonstrated the dynamic behaviour of these subtelomeres indicative for frequent exchanges between both chromosome ends (31–34).

By studying a population of >200 individuals, we obtained detailed information on the size, distribution and behaviour of these subtelomeric repeat arrays. From this study, it becomes evident that both chromosome ends present similar characteristics with regard to their size distributions and compositions, but also display discrete differences.

We found that the median of the repeat array on chromosome 4 is 21 kb larger than that on chromosome 10. A telomere position effect, in which telomeric heterochromatin has the capacity to silence genes in a distance-dependent manner, is well-documented in yeast (reviewed in ref. 35). Similar to this telomeric gene silencing, it has been postulated that the D4Z4 repeat on chromosome 4 and its homologue on chromosome 10 prevents spreading of heterochromatin into (euchromatic) regions proximal to the repeat and that partial deletion of the chromosome 4 repeat array in FSHD undermines this spacer function (36). In this light, the length difference between both subtelomeric repeats may reflect a requirement for larger subtelomeric domains on chromosome 4 to prevent gene silencing of critical gene(s) in the region proximal to the subtelomere.

An unexpectedly high frequency of 4-type repeat arrays <38 kb was observed in this survey (6 individuals = 3%). Five of them belong to class A. Whilst residing on chromosome 4, these arrays are associated with FSHD. Several possibilities may explain this finding. First, it is possible that the array does not reside on chromosome 4. Approximately 10% of individuals carry a 4-type repeat on chromosome 10 and an equal frequency of 10-type repeats on chromosome 4 is observed, we would expect the presence of 'double exchanged' alleles (i.e. a 4-type array on chromosome 10 and a 10-type allele on chromosome 4) in 1% of individuals. Second, one of the individuals carrying a 4-type array of 30 kb was assigned to class F and therefore carries three 4-type repeat arrays. It may well be that in this individual, the potential pathogenic repeat array resides on chromosome 10. Definite chomosomal assignments of these repeat arrays can only be obtained by additional PFGE analyses on agarose-embedded DNA plugs using chromosome 4- and 10-specific probes (33). Since DNA plugs of these random individuals are not available we are unable to pursue this procedure. Thirdly, ~30% of the gene carriers are asymptomatic or have subclinical FSHD signs (37,38). Moreover, a correlation has been established between the residual repeat size and the age at onset and severity of FSHD (39,40). Individuals carrying small FSHD alleles are generally more severely affected than individuals with larger FSHD alleles and it may well be that there is an incomplete penetrance for larger FSHD alleles. Since this unselected group was not clinically evaluated, it is very possible that individuals with a repeat array of <38 kb may have subclinical characteristics or are non-penetrant. In this light, it is noteworthy that the 4-type arrays in three of the five class A individuals are of >30 kb.

We confirmed that 4-type repeat arrays on chromosome 10 occur as frequently as the reverse configuration of 10-type arrays on chromosome 4 (12 and 9%, respectively). However, the composition of these translocated arrays on both chromosomes differs significantly. Although 4-type repeats on chromosome 10 tend to be homogeneous (44% of all chromosome 10 alleles), 10-type repeats on chromosome 4 are mostly not homogeneous (only 17%) but a combination of 4- and 10-type repeats (33% of all chromosome 4 alleles). This suggests a biological difference between both repeat arrays resulting in a preference of 4-type units on chromosome 4. However, that would also predict a significant difference between classes C and G, which is not observed. Therefore, we favour a mechanistic explanation.

Most repetitive sequences, whether coding or non-coding, undergo concerted evolution. This process homogenizes repetitive sequences and is thought to be important for the maintenance of the integrity of each repeat unit. Concerted evolution is achieved by a much faster intrachromosomal homogenization than interchromosomal recombination (reviewed in ref. 41). The heterogeneity observed for the D4Z4 subtelomeric repeats suggests that these loci escape concerted evolution and evolve by rather unconstrained inter- and intrachromosomal recombination. This may imply that the putative open reading frame, *DUX4*, present in each repeat unit and for which no expression has been observed *in vivo* (42), may have lost its function during the rapid expansion of this repeat array.

Our results are indicative for a multi-modal allele size distribution on both chromosomes with three equidistant peaks at intervals of ~65 kb. For minisatellite repeats, bimodal and trimodal distributions have been observed (43). This is mostly attributed to a founder effect of two ancestral alleles that have a similar dynamic behaviour and show little inter-chromosomal interactions. Since we observe the same unequal size distribution on both chromosomes and have already demonstrated a frequent interchromosomal cross-talk, a founder effect for these loci is highly unlikely. An alternative explanation that the chromatin structure of repeat sequences may impose restrictions on repeat lengths seems more likely. Degradation of high molecular weight DNA by nuclease digestion has suggested that chromatin loops contain ~50 kb of DNA (44). We have already demonstrated in mosaic individuals of *de novo* FSHD families that usually the shortest allele rearranges to an FSHD-sized allele (34). This suggests that the unequal distribution on chromosome 4 separates the repeat arrays in a premutation domain (first peak) and a normal domain (larger than the first peak).

Previously, we demonstrated that the FSHD deletion occurs mitotically in at least 40% of cases. Moreover, we showed that the presence of supernumerary fully homologous repeat arrays is a predisposing factor (in this case, the presence of 4-type repeats on chromosome 10) (34). In this study, we have identified somatic mosaicism for one of the alleles in as many as 3% of the individuals. Mosaicism was observed for both 4-type and 10-type alleles. This high mutation frequency emphasizes the importance of mitotic recombination in subtelomeric homogenization.

In conclusion, this study provides a detailed insight into the complex dynamic characteristics of the proximal subtelomeric domains on chromosomes 4 and 10. Although these domains share many properties and may frequently interact, they also display distinct differences in size and homogeneity. Their

plasticity is emphasized by a very high somatic mutation frequency. It will be interesting to examine whether this dynamic behaviour is just a side effect of mitosis and meiosis, or whether it supports distinct subtelomeric functions.

MATERIALS AND METHODS

Control samples

Genomic DNA was isolated from peripheral blood lymphocytes from 208 unrelated individuals, which were obtained in a coded form via the Dutch Bloodbank in Leiden after informed consent. DNA was extracted essentially as described by Miller *et al.* (45)

Digestion, PFGE and Southern blotting

Five micrograms of DNA was double digested with restriction enzymes *Eco*RI–*Hin*dIII (MBI Fermentas, St Leon-Rot, Germany) or *Eco*RI–*Bln*I (Amersham, Little Chalfont, UK) according to the manufacturer's instructions. After digestion, DNA was separated by PFGE on a 0.8% agarose gel (MP agarose; Boehringer, Mannheim, Germany). The electrophoresis was performed in 0.5× TBE and was run for 20 h at 8 V/cm^2. In four identical cycles, switch times increased linearly from 1 to 16 s at the end of each cycle. A pause interval of 2% of the switch time was included. After electrophoresis, the gel was stained with ethidium bromide and blotted to a Nytran$^+$ membrane (Schleicher & Schuell, Dassel, Germany). The probes used for hybridization, p13E-11 (*D4F104S1*) (26) and 9B6A (*D4Z4*) (26), were labelled by random priming with [^{32}P]dCTP, using the megaprime DNA labelling system (Amersham). Hybridizations were performed overnight at 65°C in a buffer containing 7% SDS, 10% PEG 6000 and 50 mM Na_2HPO_4 and the blots were washed at 65°C with 2× SSC/0.1% SDS (p13E-11) or 1× SSC/0.1% SDS (9B6A). The blots were then exposed to phosphorimager screens. After exposure, the alleles were assigned to their respective chromosomes based on their *Bln*I sensitivity. Sizes were estimated independently by two individuals according to a 48 kb marker and λ concatamers.

Classification of individuals

All individuals were analysed for their allele sizes and origin of repeats and classified according to Table 1. Class A represents a standard repeat array distribution with 4-type arrays on chromosome 4 and 10-type arrays on chromosome 10. Classes B–E represent translocated 10-type arrays on chromosomes 4, either homogeneous (B and E) or heterogeneous (C and D). The same holds true for translocated homogeneous (F and I) and heterogeneous (G and H) 4-type arrays on chromosome 10. Class X displays complex repeat array configurations which could not be assigned to classes A–I. Class M is a separate class in which all mosaic cases are represented.

Statistical analyses

Standard allele sizes of males (n = 100; class A) and females (n = 58; class A), were compared using the Mann–Whitney U-test. Allele-size distributions on chromosomes 4 and 10 in these 158 individuals were analysed according to the Kolmogorov–Smirnov test. Sensoring on chromosome 4 was performed for alleles of <38 kb (FSHD range). Differences in the composition of the translocated repeats (categories B, C and D versus F, G and H; i.e. only 4-type repeats, only 10-type repeats or a combination of both) were tested against a χ^2 distribution.

In order to determine whether the peaks in the multi-modal distribution of the allele sizes were equidistant, the following procedure was applied: allele sizes on a given chromosome were assumed to follow a mixture of two or three normal distributions with identical standard deviations but different means. Three hypotheses (H) were compared.

H_2. The distribution of the allele sizes was explained as a mixture of two normal distributions (four estimated parameters: mean of first distribution, difference between the means of the two distributions, standard deviation and proportion of alleles in the first distribution).

H_{3e}. A mixture of three normal distributions with equidistant means (with one additional estimated parameter: the proportion of alleles in the second distribution).

H_{3u}. As H_{3e} but with non-equidistant means (one more estimated parameter: the distance between the means of the second and the third distribution).

A comparison between likelihoods obtained for H_{3e} and H_{3u} yields a likelihood ratio test with 1 degree of freedom and represents the evidence for the means of the three distributions being at different distances.

ACKNOWLEDGEMENTS

This study was supported by the Dutch Organisation of Scientific Research (NWO), the Prinses Beatrix Fonds, the Muscular Dystrophy Association (USA) and the Dutch FSHD Foundation. S.v.d.M. is a Marjorie Bronfman fellow of the FSH Society.

REFERENCES

1. Kipling, D. (1995) *The Telomere.* Oxford University Press, Oxford, UK.
2. Louis, E.J. (1995) The chromosome ends of *Saccharomyces cerevisiae*. *Yeast*, **11**, 1553–1573.
3. Baird, D.M. and Royle, N.J. (1997) Sequences from higher primates orthologous to the human Xp/Yp telomere junction region reveal gross rearrangements and high levels of divergence. *Hum. Mol. Genet.*, **6**, 2291–2299.
4. van Deutekom, J.C.T., Lemmers, R.J.L.F., Grewal, P.K., van Geel, M., Romberg, S., Dauwerse, H.G., Wright, T.J., Padberg, G.W., Hofker, M.H., Hewitt, J.E. *et al.* (1996) Identification of the first gene (*FRG1*) from the FSHD region on human chromosome 4q35. *Hum. Mol. Genet.*, **5**, 581–590.
5. Wilkie, A.O., Higgs, D.R., Rack, K.A., Buckle, V.J., Spurr, N.K., FischelGhodsian, N., Ceccherini, I., Brown, W.R. and Harris, P.C. (1991) Stable length polymorphism of up to 260 kb at the tip of the short arm of human chromosome 16. *Cell*, **64**, 595–606.
6. Szostak, J.W. and Blackburn, E.H. (1982) Cloning yeast telomeres on linear plasmid vectors. *Cell*, **29**, 245–255.
7. Chan, C.S. and Tye, B.K. (1983) Organization of DNA sequences and replication origins at yeast telomeres. *Cell*, **33**, 563–573.
8. Walmsley, R.W., Chan, C.S., Tye, B.K. and Petes, T.D. (1984) Unusual DNA sequences associated with the ends of yeast chromosomes. *Nature*, **310**, 157–160.
9. Horowitz, H. and Haber, J.E. (1984) Subtelomeric regions of yeast chromosomes contain a 36 base-pair tandemly repeated sequence. *Nucleic Acids Res.*, **12**, 7105–7121.

10. Horowitz, H., Thorburn, P. and Haber, J.E. (1984) Rearrangements of highly polymorphic regions near telomeres of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **4**, 2509–2517.
11. Wallace, B.M. and Hulten, M.A. (1985) Meiotic chromosome pairing in the normal human female. *Ann. Hum. Genet.*, **49**, 215–226.
12. Donis-Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T.P., Bowden, D.W., Smith, D.R., Lander, E.S. *et al.* (1987) A genetic linkage map of the human genome. *Cell*, **51**, 319–337.
13. Zakian, V.A. (1989) Structure and function of telomeres. *Annu. Rev. Genet.*, **23**, 579–604.
14. NIH/CEPH Collaborative Mapping Group (1992) A comprehensive genetic linkage map of the human genome. *Science*, **258**, 67–86.
15. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. and Lathrop, M. (1992) A second-generation linkage map of the human genome. *Nature*, **359**, 794–801.
16. Flint, J., Bates, G.P., Clark, K., Dorman, A., Willingham, D., Roe, B.A., Micklem, G., Higgs, D.R. and Louis, E.J. (1997) Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. *Hum. Mol. Genet.*, **6**, 1305–1314.
17. Carlson, M., Celenza, J.L. and Eng, F.J. (1985) Evolution of the dispersed SUC gene family of *Saccharomyces* by rearrangements of chromosome telomeres. *Mol. Cell. Biol.*, **5**, 2894–2902.
18. Horowitz, H. and Haber, J.E. (1985) Identification of autonomously replicating circular subtelomeric Y′ elements in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **5**, 2369–2380.
19. Weiner, A.M., Deininger, P.L. and Efstratiadis, A. (1986) Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.*, **55**, 631–661.
20. Gottschling, D.E., Aparicio, O.M., Billington, B.L. and Zakian, V.A. (1990) Position effect at *S.cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell*, **63**, 751–762.
21. Biessmann, H. and Mason, J.M. (1992) Genetics and molecular biology of telomeres. *Adv. Genet.*, **30**, 185–249.
22. Lamb, J., Wilkie, A.O., Harris, P.C., Buckle, V.J., Lindenbaum, R.H., Barton, N.J., Reeders, S.T., Weatherall, D.J. and Higgs, D.R. (1989) Detection of breakpoints in submicroscopic chromosomal translocation, illustrating an important mechanism for genetic disease. *Lancet*, **2**, 819–824.
23. Altherr, M.R., Bengtsson, U., Elder, F.F., Ledbetter, D.H., Wasmuth, J.J., McDonald, M.E., Gusella, J.F. and Greenberg, F. (1991) Molecular confirmation of Wolf–Hirschhorn syndrome with a subtle translocation of chromosome 4. *Am. J. Hum. Genet.*, **49**, 1235–1242.
24. Kuwano, A., Ledbetter, S.A., Dobyns, W.B., Emanuel, B.S. and Ledbetter, D.H. (1991) Detection of deletions and cryptic translocations in Miller–Dieker syndrome by *in situ* hybridization. *Am. J. Hum. Genet.*, **49**, 707–714.
25. Flint, J., Wilkie, A.O.M., Buckle, V.J., Winter, R.M., Holland, A.J. and McDermid, H.E. (1995) The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nature Genet.*, **9**, 132–140.
26. Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R. *et al.* (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nature Genet.*, **2**, 26–30.
27. van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H. and Frants, R.R. (1993) FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, **2**, 2037–2042.
28. Bakker, E., Wijmenga, C., Vossen, R.H., Padberg, G.W., Hewitt, J., van der Wielen, M., Rasmussen, K. and Frants, R.R. (1995) The FSHD-linked locus *D4F104S1* (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve*, **2**, S39–S44.
29. Deidda, G., Cacurri, S., Grisanti, P., Vigneti, E., Piazzo, N. and Felicetti, L. (1995) Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the Facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet.*, **3**, 155–167.
30. Deidda, G., Cacurri, S., Piazzo, N. and Felicetti, L. (1996) Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). *J. Med. Genet.*, **33**, 361–365.
31. van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W. and Frants, R.R. (1996) Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum. Mol. Genet.*, **5**, 1997–2003.
32. Cacurri, S., Piazzo, N., Deidda, G., Vigneti, E., Galluzzi, G., Colantoni, L., Merico, B., Ricci, E. and Felicetti, L. (1998) Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric *Kpn*I repeat units implicated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **63**, 181–190.
33. Lemmers, R.J.L.F., van der Maarel, S.M., van Deutekom, J.C.T., van der Wielen, M.J.R., Deidda, G., Dauwerse, H.G., Hewitt, J., Hofker, M., Bakker, E., Padberg, G.W. *et al.* (1998) Inter- and intrachromosomal subtelomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum. Mol. Genet.*, **7**, 1207–1214.
34. van der Maarel, S.M., Deidda, G., Lemmers, R.J., van Overveld, P.G., van der Wielen, M., Hewitt, J.E., Sandkuijl, L., Bakker, B., van Ommen, G.J., Padberg, G.W. *et al.* (2000) *De novo* facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am. J. Hum. Genet.*, **66**, 26–35.
35. Guarente, L. (1999) Diverse and dynamic functions of the Sir silencing complex. *Nature Genet.*, **23**, 281–285.
36. Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., Altherr, M.R. *et al.* (1994) The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res.*, **2**, 225–234.
37. Padberg, G.W. (1982) *Facioscapulohumeral disease*. Thesis, Leiden University, Leiden, The Netherlands.
38. Padberg, G.W., Lunt, P.W., Koch, M. and Fardeau, M. (1991) Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromusc. Disord.*, **1**, 231–234.
39. Lunt, P.W., Jardine, P.E., Koch, M.C., Maynard, J., Osborn, M., Williams, M., Harper, P.S. and Upadhyaya, M. (1995) Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum. Mol. Genet.*, **4**, 951–958.
40. Tawil, R., Forrester, J., Griggs, R.C., Mendell, J., Kissel, J., McDermott, M., King, W., Weiffenbach, B. and Figlewicz, D. (1996) Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Ann. Neurol.*, **39**, 744–748.
41. Liao, D. (1999) Concerted evolution: molecular mechanism and biological implications. *Am. J. Hum. Genet.*, **64**, 24–30.
42. Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. *et al.* (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, **236**, 25–32.
43. Stead, J.D. and Jeffreys, A.J. (2000) Allele diversity and germline mutation at the insulin minisatellite. *Hum. Mol. Genet.*, **9**, 713–723.
44. Filipski, J., Leblanc, J., Youdale, T., Sikorska, M. and Walker, P.R. (1990) Periodicity of DNA folding in higher order chromatin structures. *EMBO J.*, **9**, 1319–1327.
45. Miller, S.A., Dykes, D.D. and Polesky, H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, **16**, 1215.