The DNA replication FoSTeS/MMBIR mechanism can generate human genomic, genic, and exonic complex rearrangements

Feng Zhang[1], Mehrdad Khajavi[1], Anne M. Connolly[4], Charles F. Towne[5], Sat Dev Batish[5], James R. Lupski[1,2,3]

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, 77030, USA. [2]Department of Pediatrics, Baylor College of Medicine, Houston, TX, 77030, USA. [3]Texas Children's Hospital, Houston, TX, 77030, USA. [4]Department of Neurology, Washington University School of Medicine, Saint Louis, MO 63110, USA. [5]Athena Diagnostics, Inc., Worcester, MA 01605, USA

Correspondence should be addressed to Dr. James R. Lupski (jlupski@bcm.tmc.edu)

Nature Genetics: doi:10.1038/ng.399

Supplementary Information

Supplementary Tables 1 to 4

Supplementary Figs. 1 to 10 with Legends

References

Nature Genetics: doi:10.1038/ng.399

Supplementary Table 1. Characteristics of the simple nonrecurrent rearrangements in 17p

Scale	Sample	Rearrangement	Size	Microhomology
Genomic	2661	dup	11.2 Mb	AT
	1861	dup	7.7 Mb	N.A.
	2488	dup	7.5 Mb	N.A.
	2362	dup	7.0 Mb	N.A.
	2440	dup	4.9 Mb	N.A.
Exonic	A10	del	16.8 kb	GATT
	A21	del	11.6 kb	C
	A11	del	7.2 kb	AC
	A14	del	212 bp	GC
	A12	del	9 bp	GACG

Abbreviation: del, deletion; dup, duplication; N.A., not available.

Supplementary Table 2. Five out of 23 (22%) CNV breakpoints sequenced in Perry *et al.*[1] show sequence complexities consistent with two or more FoSTeS/MMBIR events

Example	Genomic location	FoSTeS events	Microhomology
1	2q21.1	×2	GAGG, CA
2	2q32.3	×4	GAAGT, AATTC, ATTT, AAAACA
3	3q22.1	×2	CT, TG
4	4q26	×2	TTA, GAATA
5	5q33.1	×2	CAA, TCT

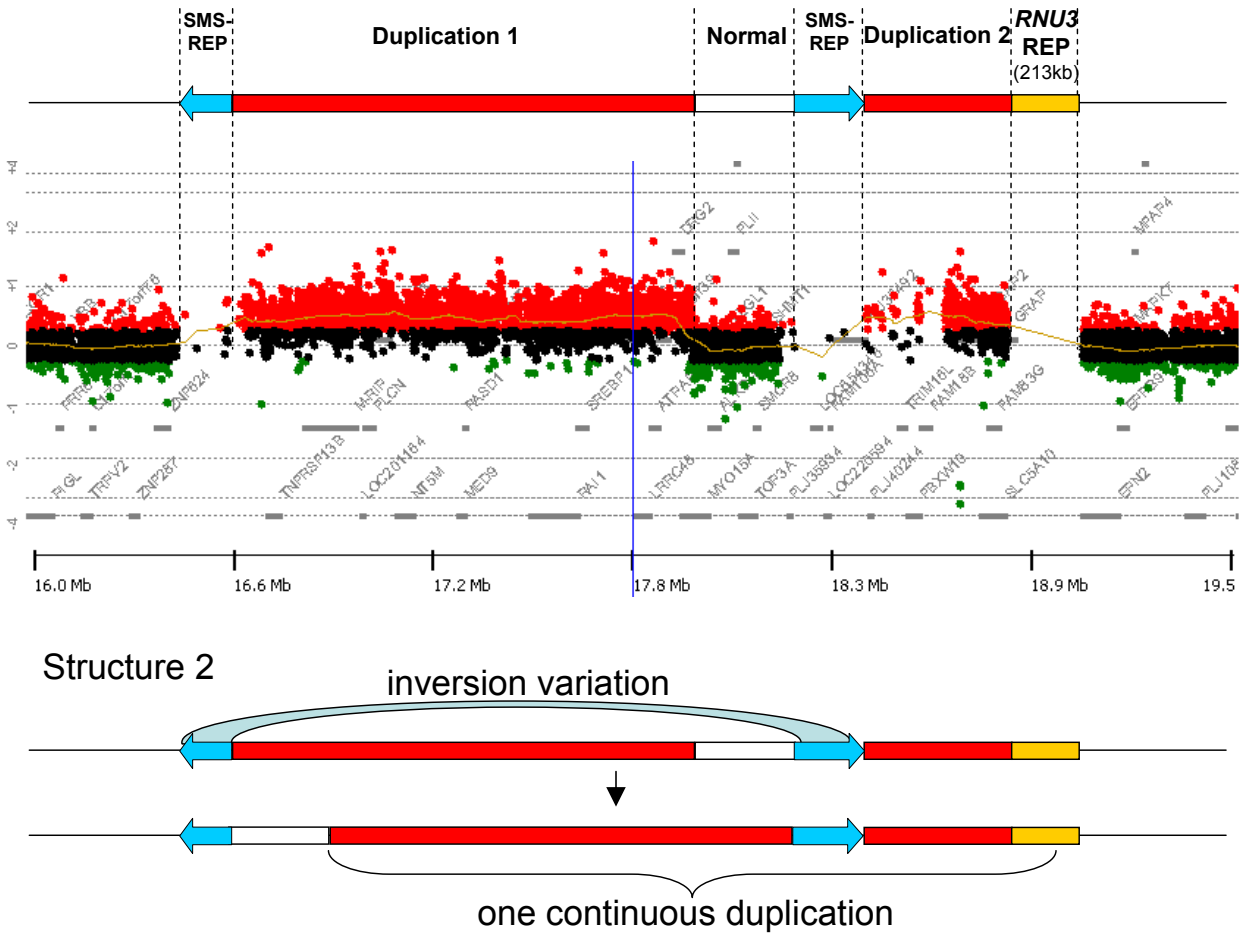
Supplementary Table 3. Reported complex gene rearrangements consistent with multiple FoSTeS/MMBIR events (http://www.hgmd.cf.ac.uk/ac/index.php)

Gene	Genomic location	Exon(s) involved	Events	Microhomology	Reference
*SLC3A1**	2p21	ex 2	×2	CTT, AAA	2
ATP2C1	3q22.1	ex 20	×2	AG, TGTGT	3
FOXL2	3q22.3	ex 1	×2	GC, GCAGCCGCGG	4
DSPP	4q22.1	ex 5	×2	AGCAG, AGCAGTGACAGCA	5
HPS1	10q24.2	ex 12-20	×2	GTTACTGGG, CCCCTCAGAGCC	6
HNF1B	17q12	ex 5	×2	TATT, AACA	7
BRCA1	17q21.31	ex 21, 22	×2	CCT, CACTGCACTCCA-GCCTGGGTGACAGAG	8
*SLC5A5**	19p13.11	ex 3-6	×2	CATC, AC	9
PRPF31	19q13.42	ex 1	×2	AAC, TCCTGGG	10
*OFD1**	Xp22.2	ex 7-9	×2	TCA, CTCAA	11
IL2RG	Xq13.1	ex 4	×2	TG, GTGA	12
*GLA**	Xq22.1	ex 5-7	×2	AC, AT	13
*F9**	Xq27.1	ex 4	×2	ATT, CAAAAA	14
*F8**	Xq28	ex 16	×2	AACT, TCA	15
*IDS**	Xq28	ex 5-7	×4	AATACATTT, GA, CAAGAG, CA	16
L1CAM	Xq28	ex 2	×2	GCT, GC	17
*MECP2**	Xq28	ex 4	×2	CC, AGCCC	18

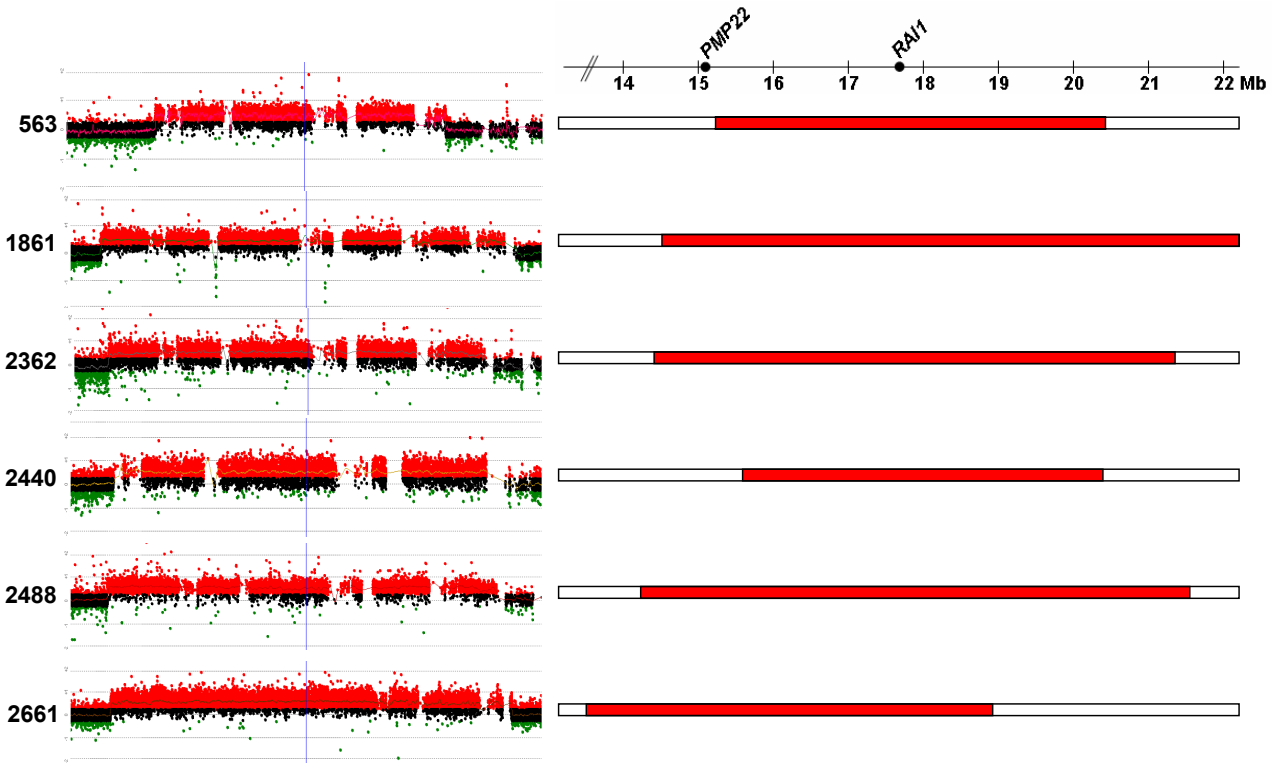
*Some complex gene rearrangements have been hypothesized to be caused by "serial replication slippage"[19,20], an alternative replication-based model proposed for small complex rearrangements.

Supplementary Table 4. Primers of achieved long-range PCR amplifications and the breakpoint junction products for sequencing in a proportion of nonrecurrent rearrangements on 17p

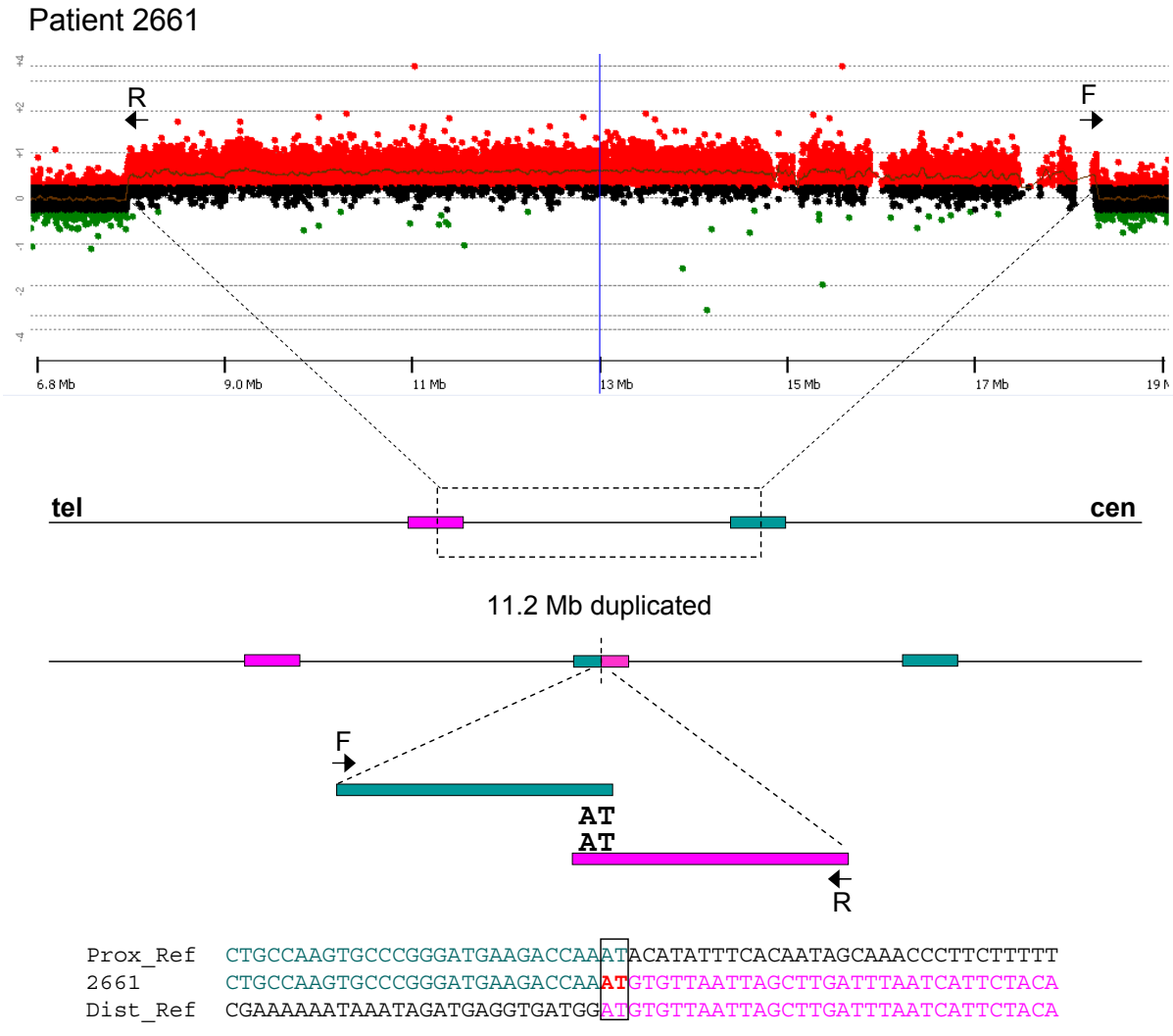
Sample	Primer coordinate (+/- strand), UCSC hg18		Size of product for Sequencing (kb)
	Forward	Reverse	
563	chr17:20775964-20775983 (+)	chr17:15225900-15225920 (-)	1.3
621	chr17:17034792-17034812 (+)	chr17:17021067-17021090 (-)	0.5
	chr17:17033028-17033053 (+)	Same with the forward primer	10
1229	chr17:15015801-15015820 (+)	chr17:15054771-15054791 (-)	0.5
2661	chr17:19124269-19124292 (+)	chr17:7977899-7977918 (-)	4.2
2695	chr17:16174939-16174959 (+)	chr17:16137968-16137993 (-)	3.9
2711	chr17:10870486-10870509 (+)	chr17:10955976-10955998 (-)	2.0
A10	chr17:15078686-15078705 (+)	chr17:15096288-15096307 (-)	0.8
A11	chr17:15096041-15096060 (+)	chr17:15104032-15104052 (-)	0.8
A12	chr17:15104214-15104233 (+)	chr17:15105252-15105272 (-)	1.1
A14	chr17:15104214-15104233 (+)	chr17:15105252-15105272 (-)	0.9
A15	chr17:15079113-15079132 (+)	chr17:15088798-15088817 (-)	1.1
A21	chr17:15093195-15093218 (+)	chr17:15105743-15105764 (-)	1.0



Supplementary Figure 1. Rearrangement for patient 2543. Oligonucleotide aCGH revealed a complex rearrangement with two separate duplications (structure 1) in patient 2543. Since NAHR between the SMS repeats (distal and middle SMS-REPs in reverse orientation) can lead to an inversion allele and such an inversion variation has been reported[21], the complexity in patient 2543 can alternatively be caused by a single duplication that occurred on an inversion allele (structure 2).

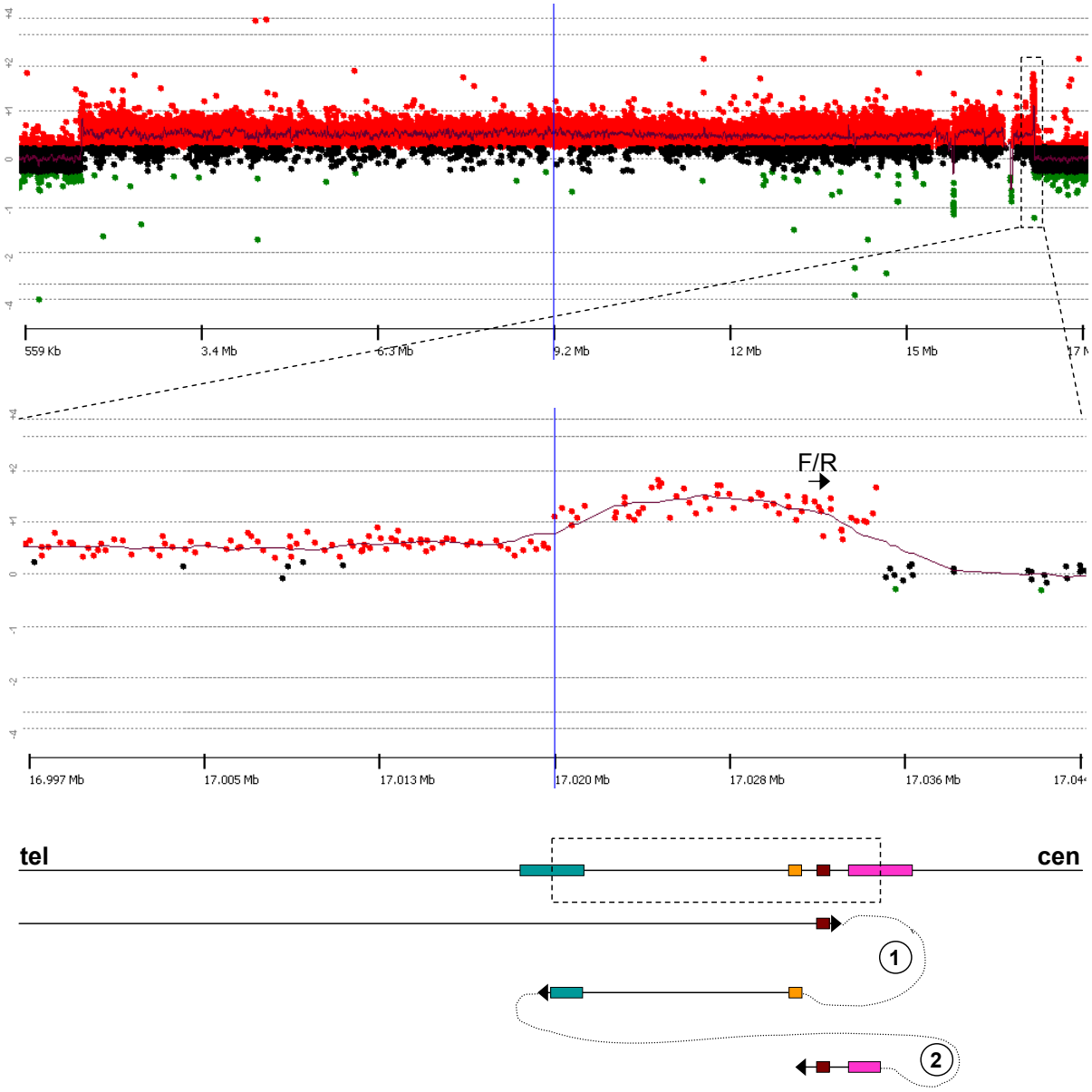


Supplementary Figure 2. No sequence complexity was found in 6 out of 14 nonrecurrent 17p11.2 duplications based on oligonucleotide aCGH. Further breakpoint sequence analyses revealed the FoSTeS/MMBIR-mediated complexity in patient 563.



Supplementary Figure 3. Junction sequence analysis of the duplication identified in 2661. Two base-pairs of microhomology (AT) were found at the breakpoint.

Patient 621



FoSTeS 1

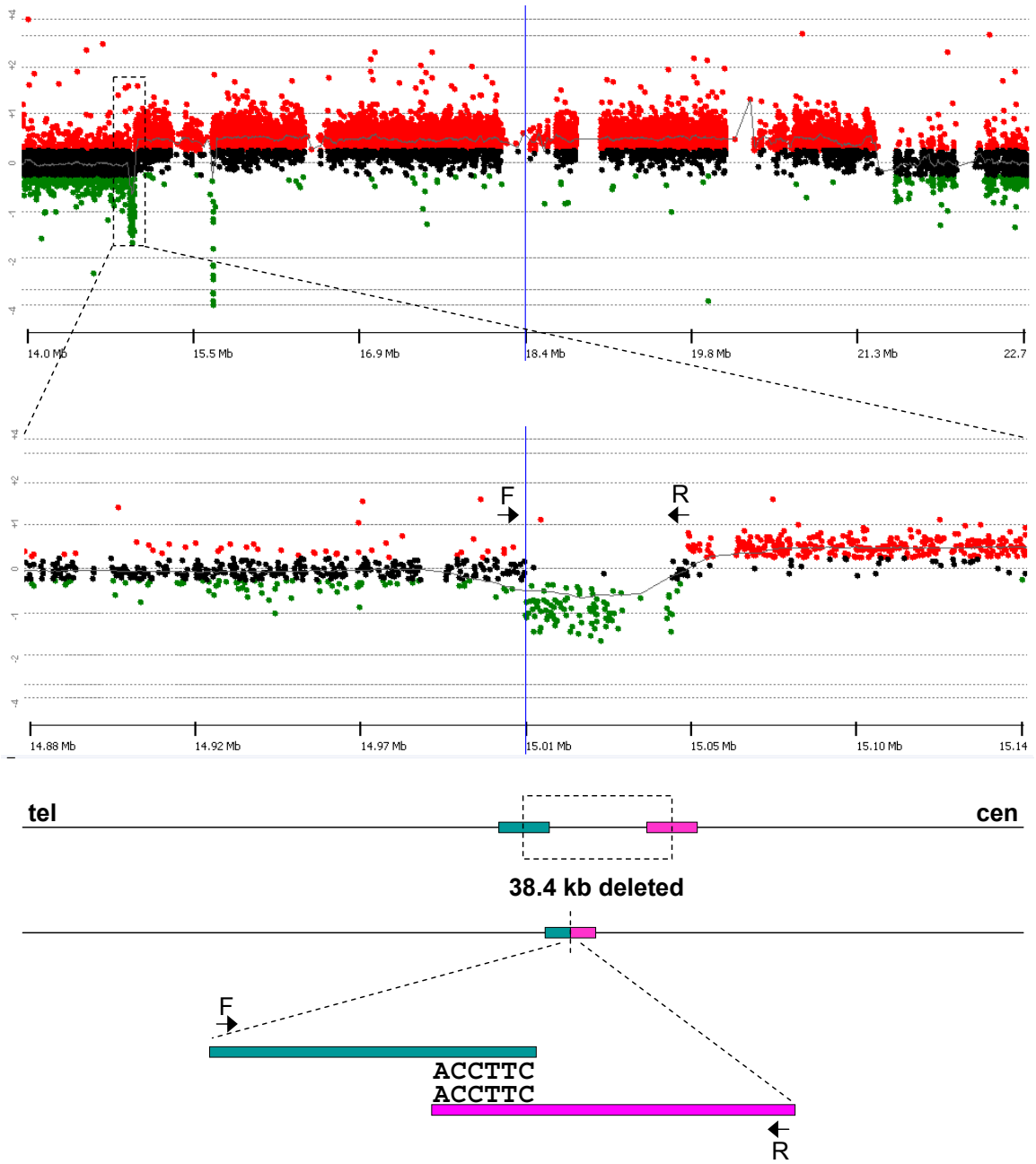
```
ProxRef1+ CAGGATGGTCTCGATTTCCTGACCTCGTGATCCGCCTGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGC
621_1+    CAGGATGGTCTCGATTTCCTGACCTCGTGATCCGCCAGCCAGGGGTTTATACTCCCTATCTTAATGGGAT
DistRef2- CCACCTTGCTCCTAAAGTAGCAGAACCTGTTCCACCAGCCAGGGGTTTATACTCCCTATCTTAATGGGAT
```

FoSTeS 2

```
ProxRef2+ ATGTTTTTAATGACTGCTGCATCTTTGTAAAACGTTTGGTCATCTAACAGATGGTTTTAAAGTGTACAAT
621_2-    ATGTTTTTAATGACTGCTGCATCTTTGTAAAACGTTTGGTGGTCACTCAGTCACCTAGGGAATCTAGGAC
DistRef2+ AGGGACTCCTGGAGTCAGGGGAGATGACAGAATCATTGGTGGTCACTCAGTCACCTAGGGAATCTAGGAC
```

Supplementary Figure 4. Sequence analysis of one breakpoint junction of the complex rearrangement in patient 621. An about 10 kb junction product was amplified by only one primer at the proximal end (with respect to the reference sequence) of the complex rearrangement region. Sequencing analysis revealed two of the FoSTeS/MMBIR events causing complexity.

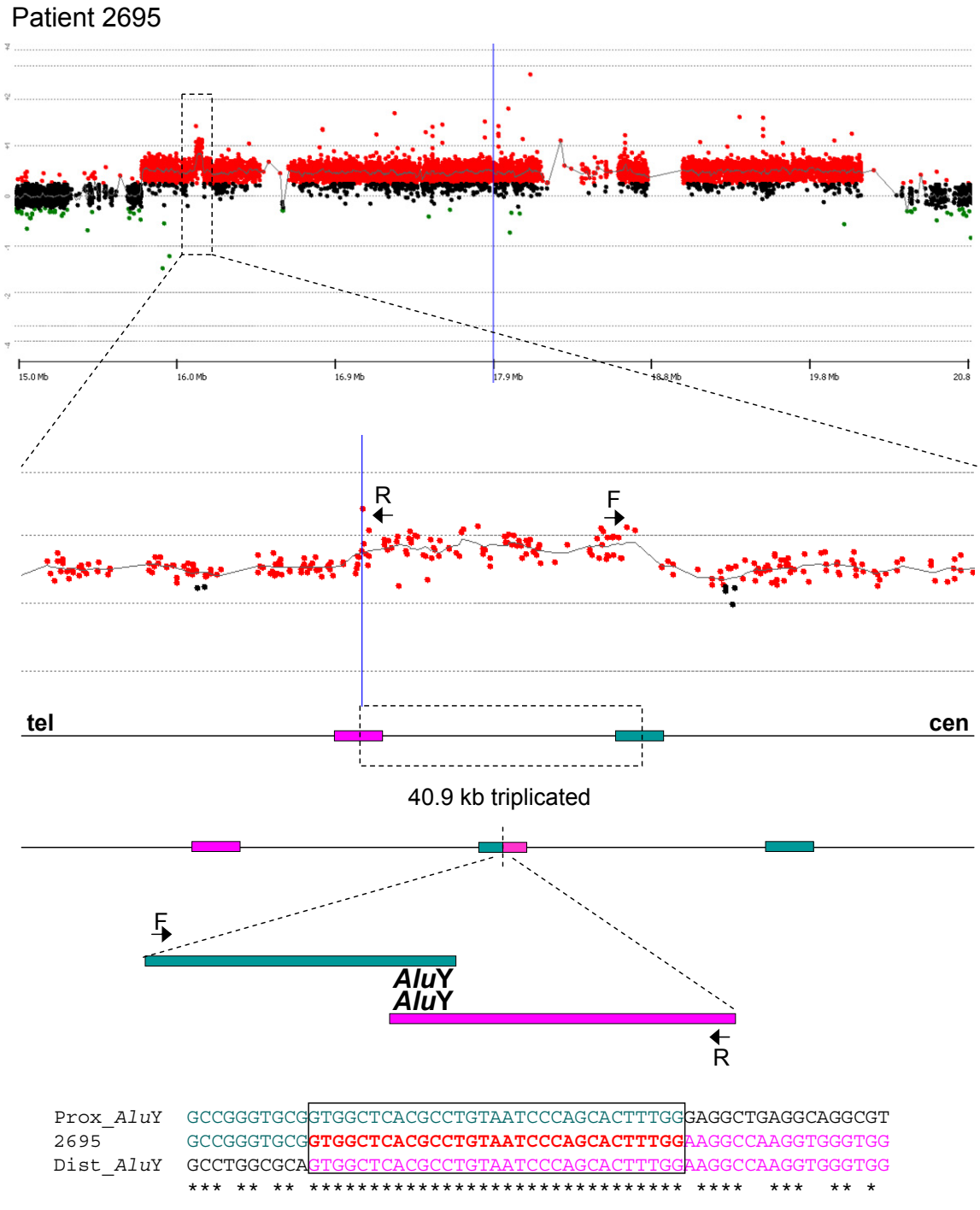
Patient 1229



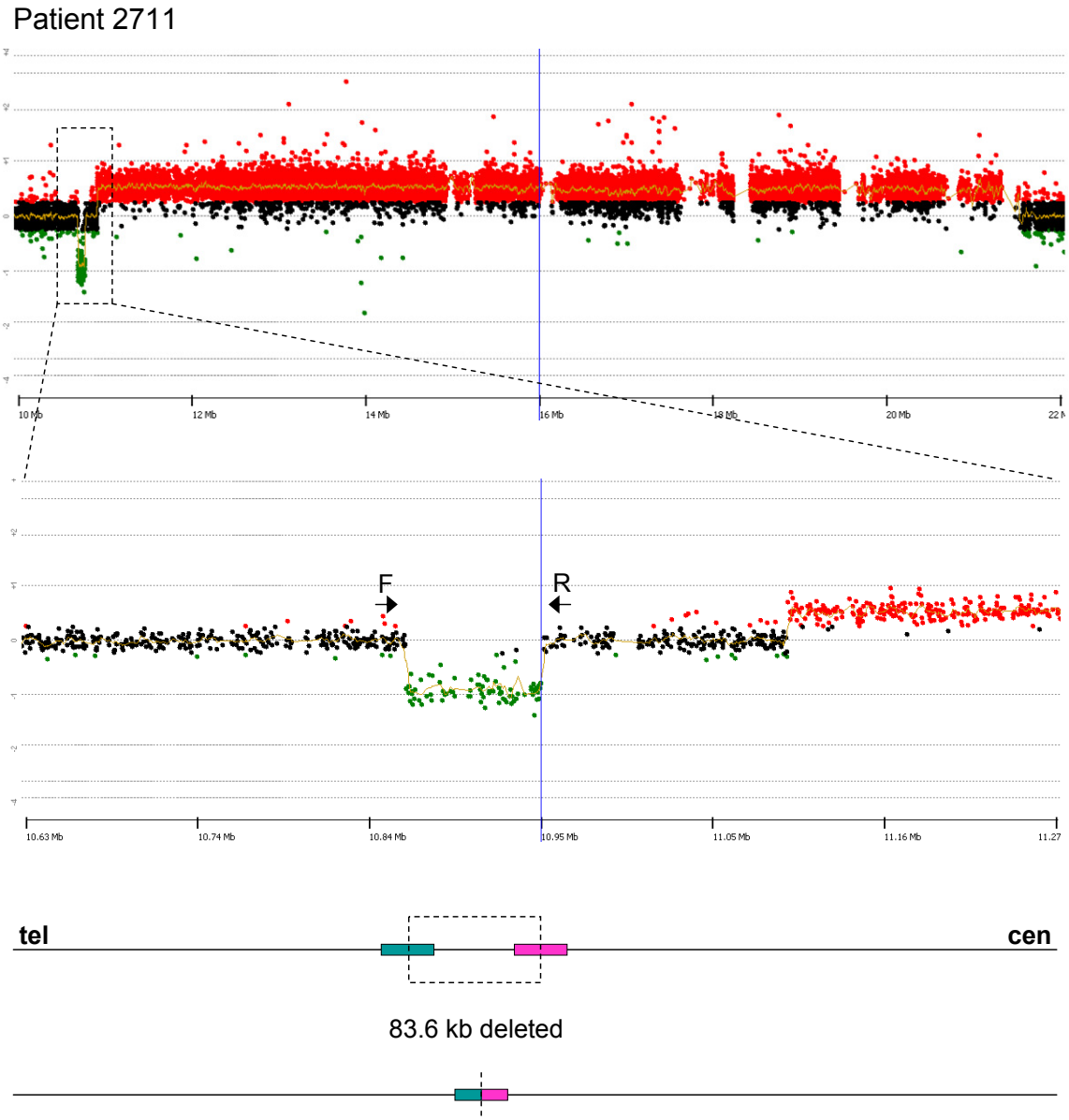
FoSTeS 1

```
Dist_Ref CTCCACATGTGGAATCTCCCCAGACCCCATGACCTTCATACACATTTCACCATTTCCCAGAGCCAGATTT
1229     CTCCACATGTGGAATCTCCCCAGACCCCATGACCTTCTGCTTCCTAACAACCACTGAGATAATTTGATAT
Prox_Ref TTACTGCAGGGGATTGGAGTCAACATAATTAACCTTCTGCTTCCTAACAACCACTGAGATAATTTGATAT
```

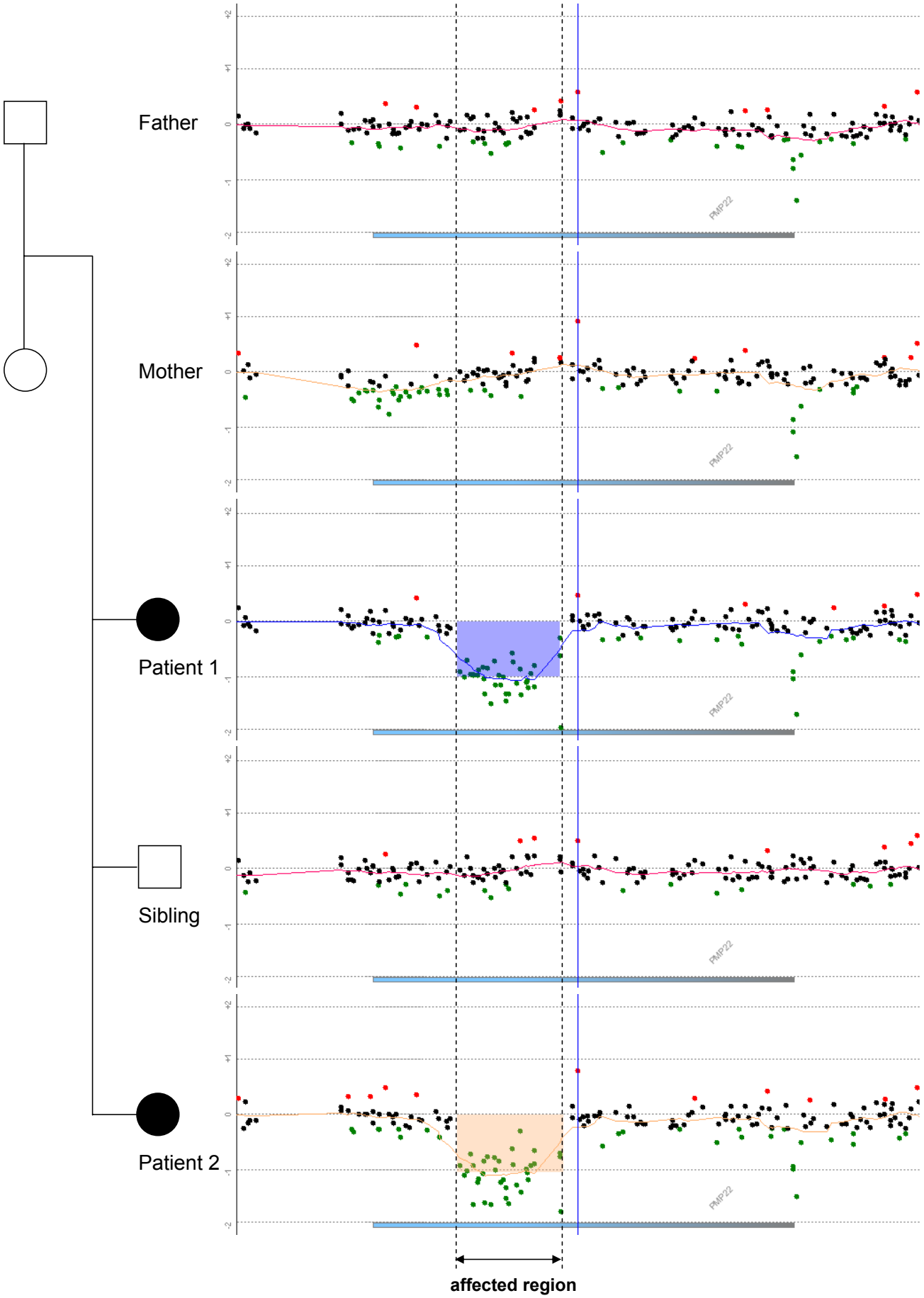
Supplementary Figure 5. Sequence analysis of one junction of a complex rearrangement identified in patient 1229. At this deletion breakpoint, a microhomology of 6 bp (ACCTTC) was found.



Supplementary Figure 6. Junction sequence analysis of the triplicated region in the complex rearrangement in patient 2695. At the breakpoint, there is 31 bp microhomology between two *Alu*Y elements.



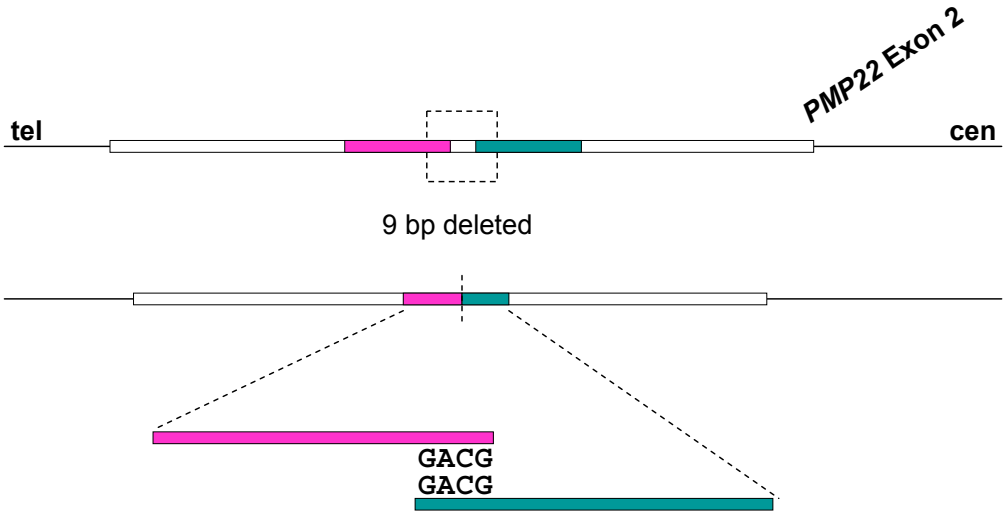
Supplementary Figure 7. Sequence analysis of one junction of the complex rearrangement in patient 2711. A “C” nucleotide was inserted at the deletion breakpoint.



Supplementary Figure 8. Oligonucleotide aCGH analysis of the *PMP22* gene region in one CMT1 family. The identical small deletions were identified in two affected siblings. No obvious copy number loss of the deletion region was found in healthy sibling and parents.

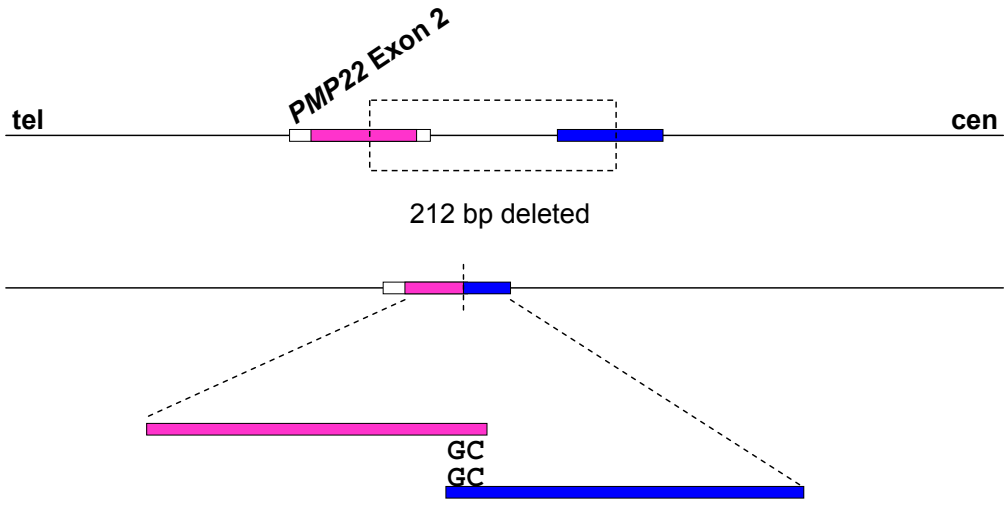
Patient A12

```
Ref    TCGTGGAGACGAACAGCAGCACCAGCACCGCGACGTGGAGGACGATGATACTCAGCAACAGGAGGAGCAT
A12    TCGTGGAGACGAACAGCAGCACCAGCACCGCGACG---------ATGATACTCAGCAACAGGAGGAGCAT
```



Patient A14

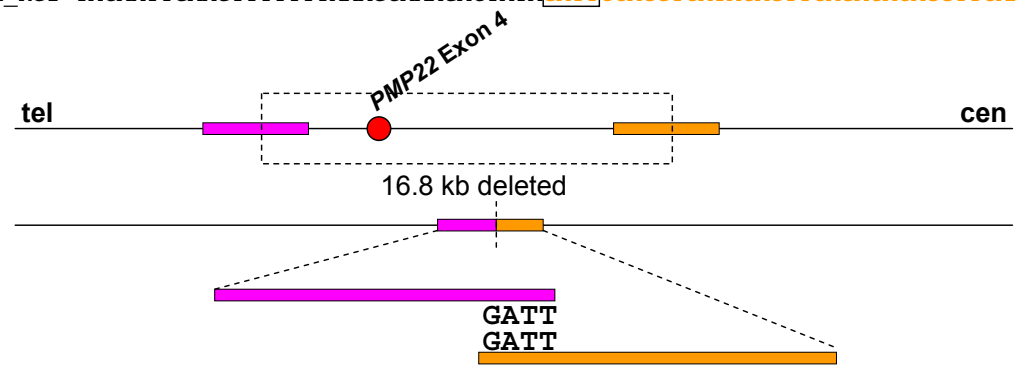
```
Dist_Ref GCACTCACGCTGACGATCGTGGAGACGAACAGCAGCACCAGCACCGCGACGTGGAGGACGATGATACTCA
A14      GCACTCACGCTGACGATCGTGGAGACGAACAGCAGCGCGGGTCAGGAGCCTTCGCGCCGCCTGCCGCCGA
Prox_Ref CCGGCCTGGCCCAGCGCCCGCAGCCCGACCGCCCGCGCGGGTCAGGAGCCTTCGCGCCGCCTGCCGCCGA
```



Supplementary Figure 9. Junction sequence analysis of small deletions within *PMP22* gene in patient A12 and A14. *PMP22* Exon 2 was represented by a white bar.

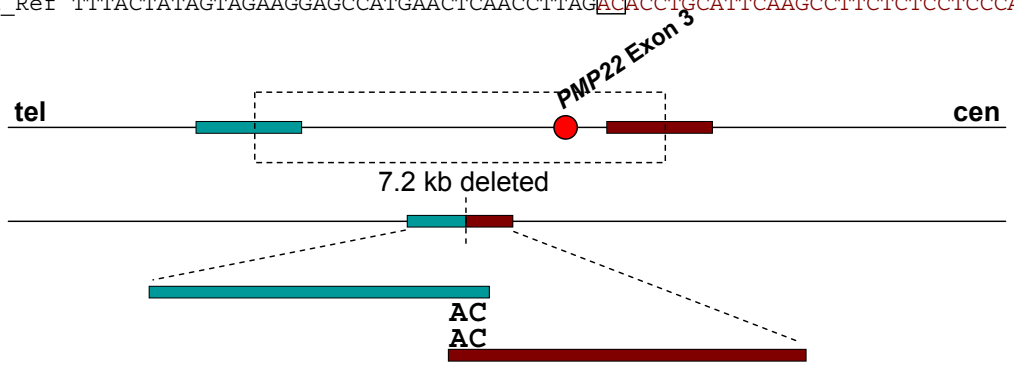
Patient A10

```
Dist_Ref CAATCCTCTCACCTCAGTCTTTCAAAGTGTTGGGATTACAGGCATTCGCCACCACACCTGGCCCAAATCC
A10_1    CAATCCTCTCACCTCAGTCTTTCAAAGTGTTGGGATTCCAGGTCATACAGTTCAGACACAGGTTCAAACA
Prox_Ref TACAATTGAAGTTTTTTTAAAGCAAAGACTATAGATTCCAGGTCATACAGTTCAGACACAGGTTCAAACA
```



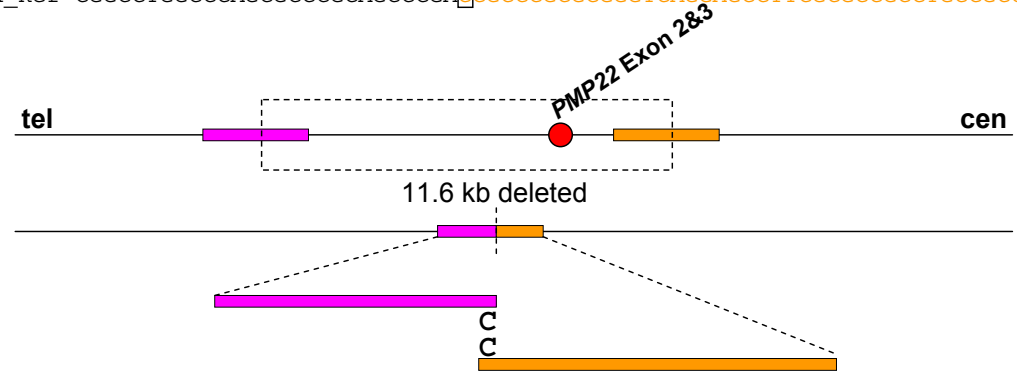
Patient A11

```
Dist_Ref GGAAATTGTCTGGTGTCCTTGCCAAAATGACAGTGTCACCAAATTGATAAACAAGTCTGGGATCGTTTTG
A11_1    GGAAATTGTCTGGTGTCCTTGCCAAAATGACAGTGTCACACCTGCATTCAAGCCTTCTCTCCTCCCAGCA
Prox_Ref TTTACTATAGTAGAAGGAGCCATGAACTCAACCTTAGACACCTGCATTCAAGCCTTCTCTCCTCCCAGCA
```



Patient A21

```
Dist_Ref ACTTCAGGTCCTACCTCAATGCCTGCTCTGCTTTTAGGAAACCAGGGTACCTGCCCCAGGTTACCTAATT
A21_1    ACTTCAGGTCCTACCTCAATGCCTGCTCCGCCCGCGCGGGTCAGGAGCCTTCGCGCCGCCTGCCGCCGAG
Prox_Ref CGGCCTGGCCCAGCGCCCGCAGCCCGACCGCCCGCGCGGGTCAGGAGCCTTCGCGCCGCCTGCCGCCGAG
```



Supplementary Figure 10. Junction sequence analysis of small deletions within the *PMP22* gene in patients A10, A11, and A21.

References

1. Perry, G.H. et al. The fine-scale and complex architecture of human copy-number variation. *Am. J. Hum. Genet.* **82**, 685-95 (2008).
2. Gitomer, W.L., Reed, B.Y., Ruml, L.A., Sakhaee, K. & Pak, C.Y. Mutations in the genomic deoxyribonucleic acid for *SLC3A1* in patients with cystinuria. *J. Clin. Endocrinol. Metab.* **83**, 3688-3694 (1998).
3. Li, H., Sun, X.K. & Zhu, X.J. Four novel mutations in *ATP2C1* found in Chinese patients with Hailey-Hailey disease. *Br. J. Dermatol.* **149**, 471-474 (2003).
4. De Baere, E. et al. Spectrum of *FOXL2* gene mutations in blepharophimosis-ptosis-epicanthus inversus (BPES) families demonstrates a genotype--phenotype correlation. *Hum. Mol. Genet.* **10**, 1591-1600 (2001).
5. Dong, J., Gu, T., Jeffords, L. & MacDougall, M. Dentin phosphoprotein compound mutation in dentin sialophosphoprotein causes dentinogenesis imperfecta type III. *Am. J. Med. Genet. A* **132A**, 305-309 (2005).
6. Griffin, A.E. et al. Detection of hemizygosity in Hermansky-Pudlak syndrome by quantitative real-time PCR. *Clin. Genet.* **68**, 23-30 (2005).
7. Bellanne-Chantelot, C. et al. Large genomic rearrangements in the hepatocyte nuclear factor-1beta (*TCF2*) gene are the most frequent cause of maturity-onset diabetes of the young type 5. *Diabetes* **54**, 3126-3132 (2005).
8. Vasickova, P. et al. High occurrence of *BRCA1* intragenic rearrangements in hereditary breast and ovarian cancer syndrome in the Czech Republic. *BMC Med. Genet.* **8**, 32 (2007).
9. Kosugi, S., Okamoto, H., Tamada, A. & Sanchez-Franco, F. A novel peculiar mutation in the sodium/iodide symporter gene in spanish siblings with iodide transport defect. *J. Clin. Endocrinol. Metab.* **87**, 3830-3836 (2002).
10. Sullivan, L.S. et al. Genomic rearrangements of the *PRPF31* gene account for 2.5% of autosomal dominant retinitis pigmentosa. *Invest. Ophthalmol. Vis. Sci.* **47**, 4579-4588 (2006).
11. Morisawa, T. et al. Novel double-deletion mutations of the *OFD1* gene creating multiple novel transcripts. *Hum. Genet.* **115**, 97-103 (2004).
12. Tsuge, I., Matsuoka, H., Abe, T., Kamachi, Y. & Torii, S. Interleukin-2 receptor gamma-chain mutations in severe combined immunodeficiency with B-lymphocytes. *Eur. J. Pediatr.* **155**, 1018-1024 (1996).
13. Kornreich, R., Bishop, D.F. & Desnick, R.J. Alpha-galactosidase A gene rearrangements causing Fabry disease. Identification of short direct repeats at breakpoints in an *Alu*-rich gene. *J. Biol. Chem.* **265**, 9319-9326 (1990).
14. Solera, J., Magallon, M., Martin-Villar, J. & Coloma, A. Factor IXMadrid 2: a deletion/insertion in factor IX gene which abolishes the sequence of the donor junction at the exon IV-intron d splice site. *Am. J. Hum. Genet.* **50**, 434-437 (1992).
15. Tavassoli, K., Eigel, A. & Horst, J. A deletion/insertion leading to the generation

of a direct repeat as a result of slipped mispairing and intragenic recombination in the factor VIII gene. *Hum. Genet.* **104**, 435-437 (1999).
16. Karsten, S., Voskoboeva, E., Krasnopolskaja, X. & Bondeson, M.L. Novel type of genetic rearrangement in the iduronate-2-sulfatase (*IDS*) gene involving deletion, duplications, and inversions. *Hum. Mutat.* **14**, 471-476 (1999).
17. Tegay, D.H., Lane, A.H., Roohi, J. & Hatchwell, E. Contiguous gene deletion involving *L1CAM* and *AVPR2* causes X-linked hydrocephalus with nephrogenic diabetes insipidus. *Am. J. Med. Genet. A* **143**, 594-598 (2007).
18. Buyse, I.M. et al. Diagnostic testing for Rett syndrome by DHPLC and direct sequencing analysis of the *MECP2* gene: identification of several novel mutations and polymorphisms. *Am. J. Hum. Genet.* **67**, 1428-1436 (2000).
19. Chen, J.M., Chuzhanova, N., Stenson, P.D., Ferec, C. & Cooper, D.N. Complex gene rearrangements caused by serial replication slippage. *Hum. Mutat.* **26**, 125-134 (2005).
20. Chen, J.M., Chuzhanova, N., Stenson, P.D., Ferec, C. & Cooper, D.N. Intrachromosomal serial replication slippage in trans gives rise to diverse genomic rearrangements involving inversions. *Hum. Mutat.* **26**, 362-373 (2005).
21. Tuzun, E. *et al.* Fine-scale structural variation of the human genome. *Nat. Genet.* **37**, 727-732 (2005).