The American Journal of Human Genetics, Volume *95*

Supplemental Data

The *Alu*-Rich Genomic Architecture of *SPAST* Predisposes to Diverse and Functionally Distinct Disease-Associated CNV Alleles

Philip M. Boone, Bo Yuan, Ian M. Campbell, Jennifer C. Scull, Marjorie A. Withers, Brett C. Baggett, Christine R. Beck, Christine J. Shaw, Pawel Stankiewicz, Paolo Moretti, Wendy E. Goodwin, Nichole Hein, John K. Fink, Moon-Woo Seong, Soo Hyun Seo, Sung Sup Park, Izabela D. Karbassi, Sat Dev Batish, Andrés Ordóñez-Ugalde, Beatriz Quintáns, María-Jesús Sobrido, Susanne Stemmler, and James R. Lupski

Supplemental Figures

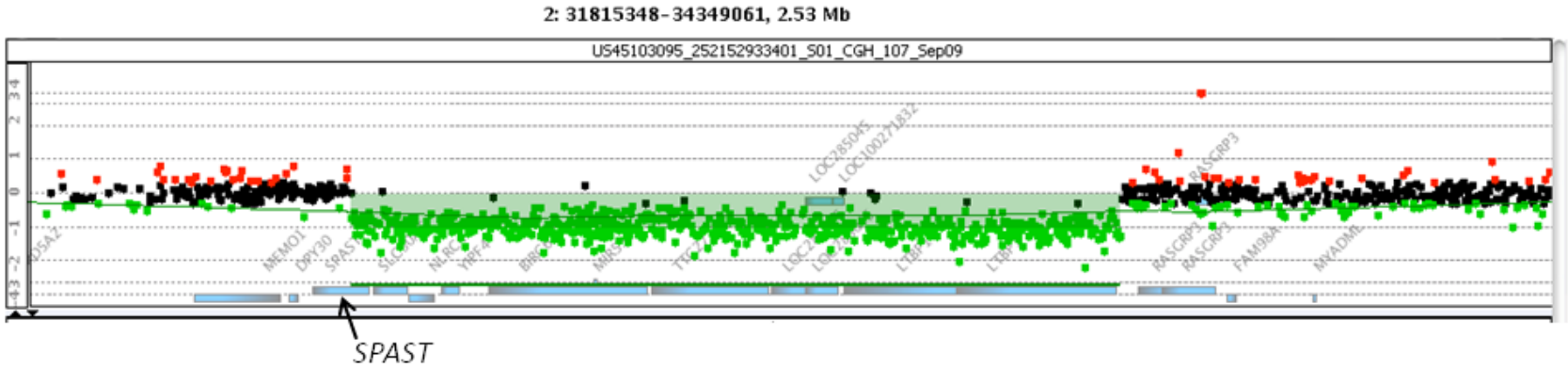


Figure S1. Array CGH result for Ruhr 29. The CNV in subject Ruhr 29 spanned beyond the boundaries of the custom array. Thus, a 1M probe genome-wide array was used to confirm and map this subject's deletion.

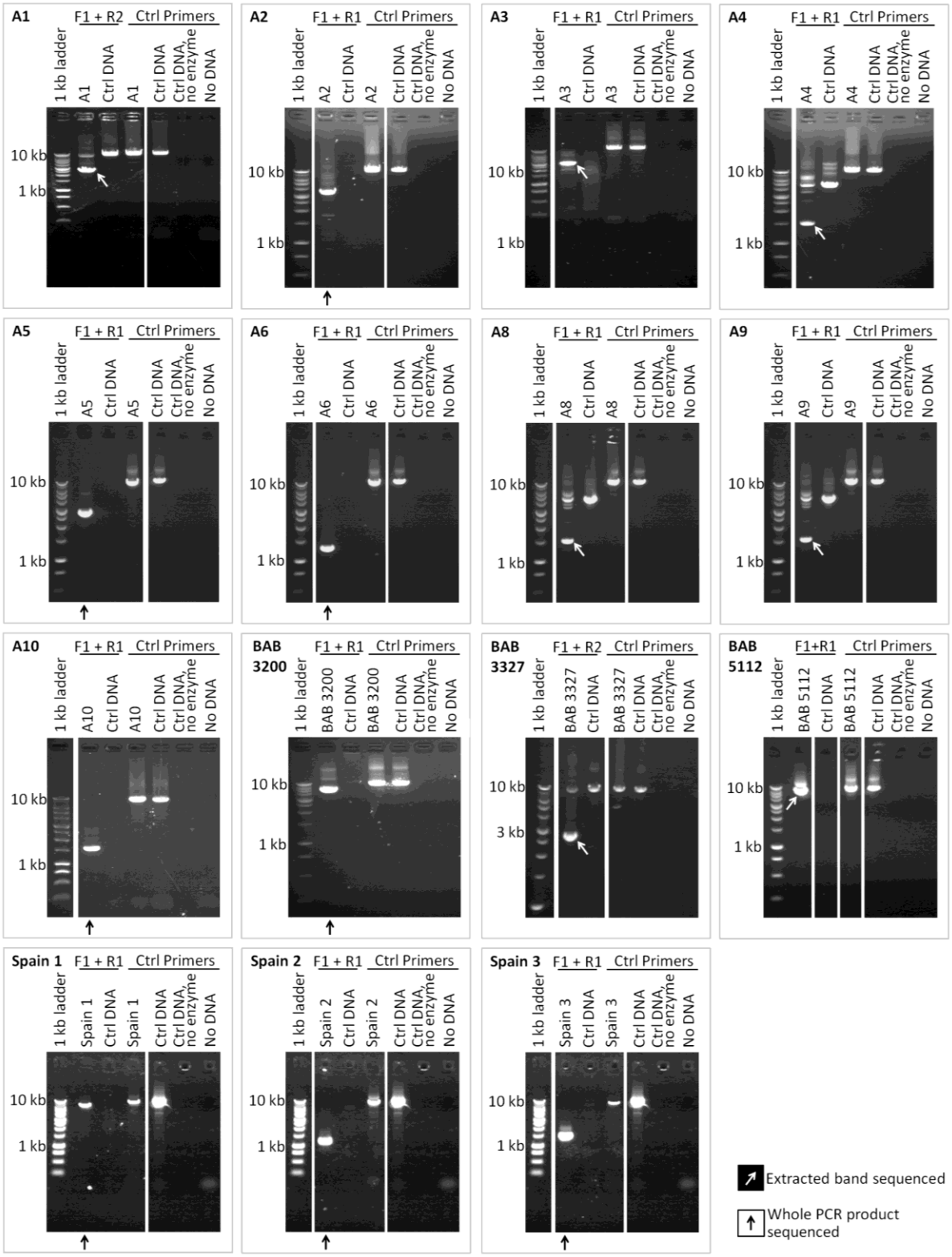


Figure S2. CNV breakpoint PCRs. Note that no PCR product was obtained for Ruhr 32. This may be a complex CNV.

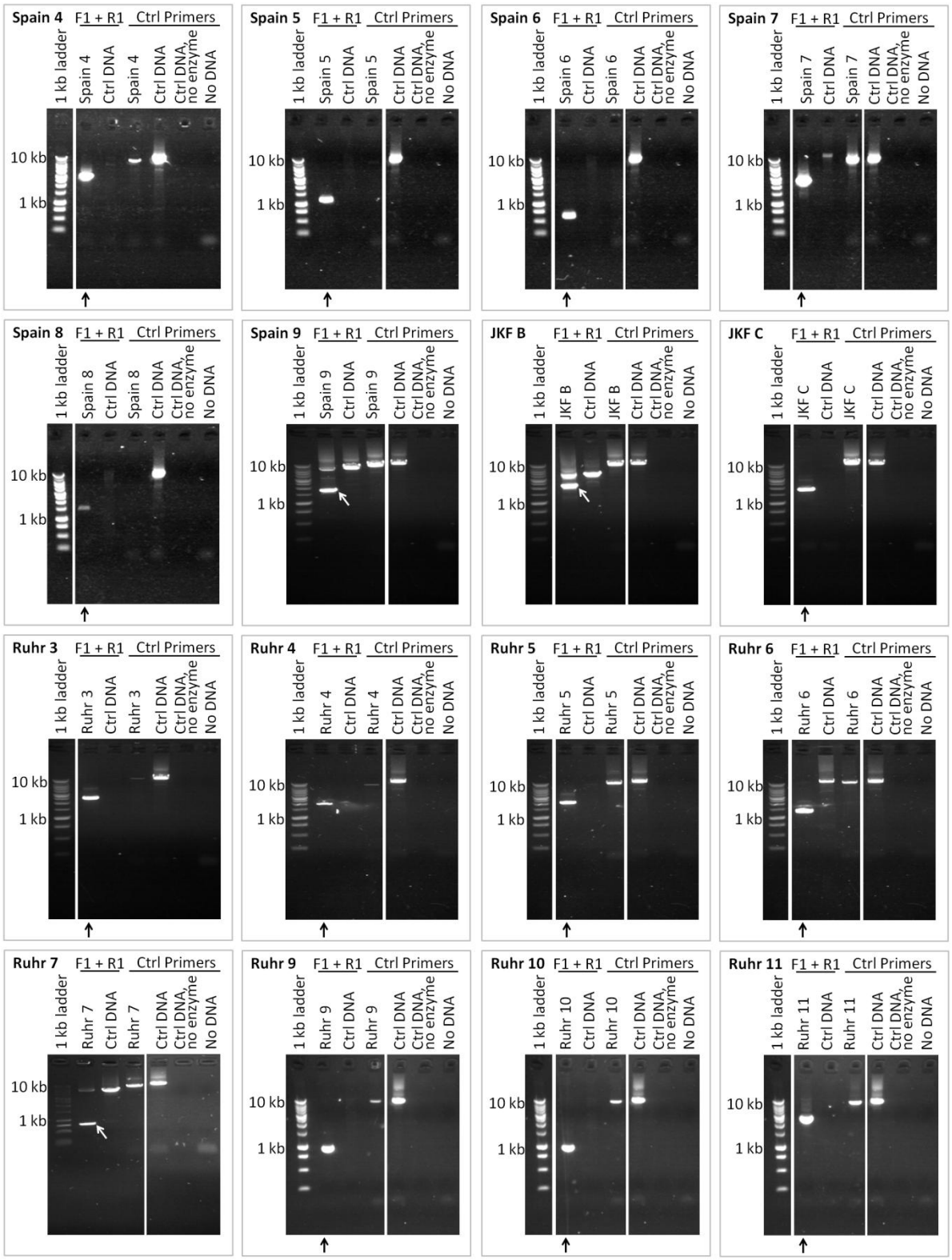


Figure S2, cont'd.

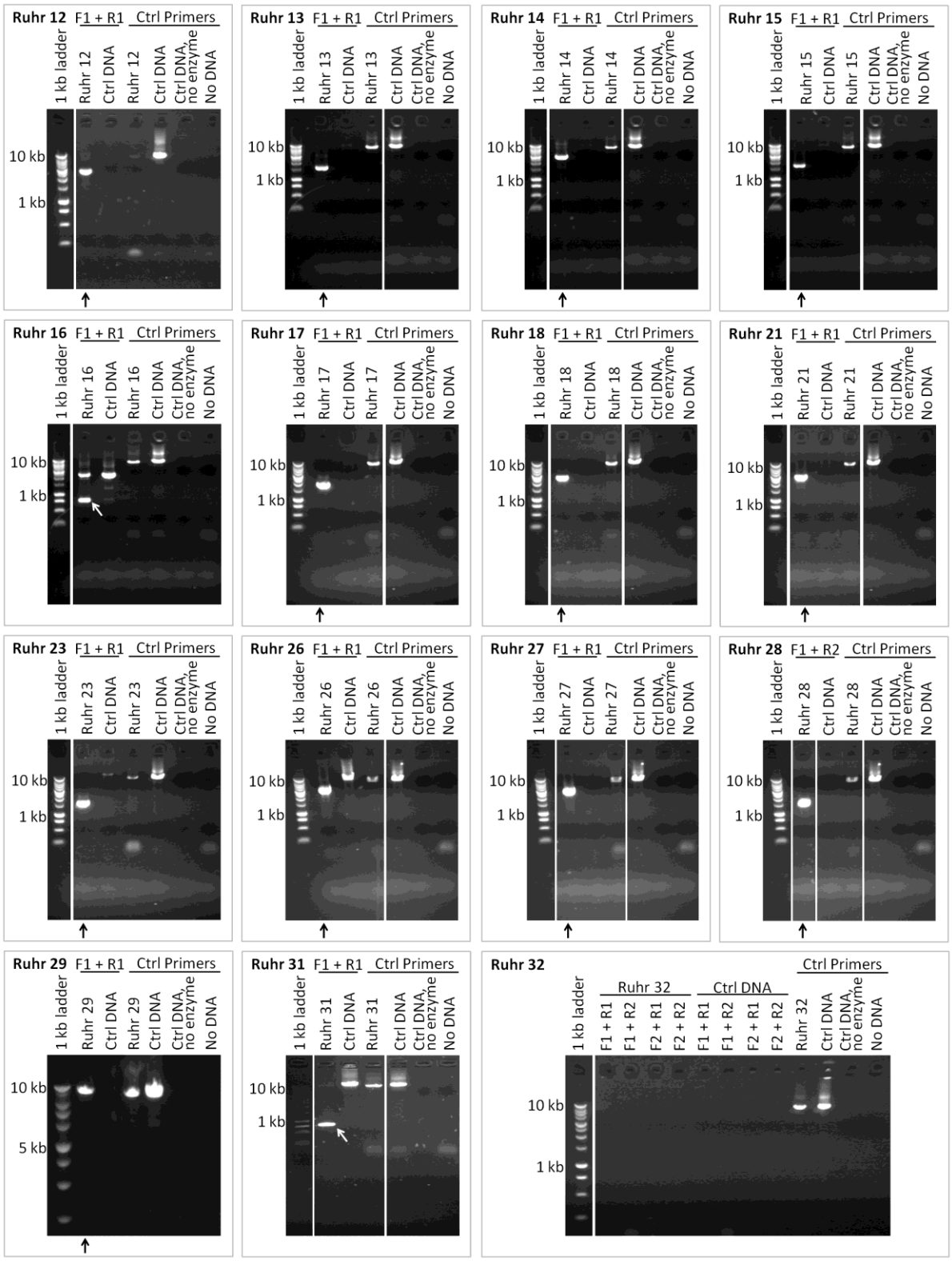


Figure S2, cont'd.

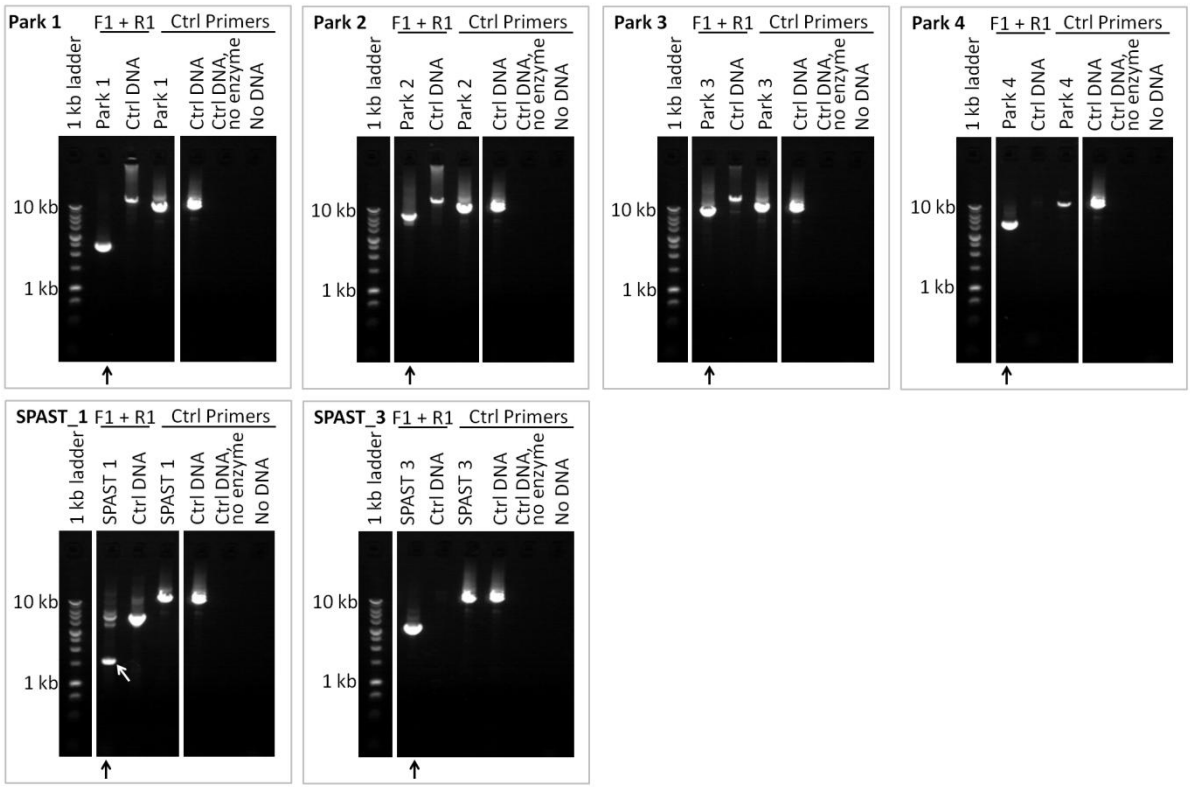


Figure S2, cont'd.

A1

Reverse complement sequence

```
Chr2:32378859 CTTTAGGCTGGGCAAAGTGGCTCACGTCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTTAGG 32378937
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||   ||||||||||   ||||||||
A1         305 CTTTAGGCTGGGCAAAGTGGCTCACGTCTGTAATCCCAGCACTTTGGGAGGCCGAGACAGGCGGATCACA--AGGTTAGG 229
                | ||||||||   |||||||||| ||||||||||||||||||||||||||||||||||| |||||||--|||| |||
chr2:32384360 TAATTGGCTGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACAGGCAGATCACA--AGGTCAGG 32384437
                rs187211683* * rs4519572                                rs4407291*             *rs4530415
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

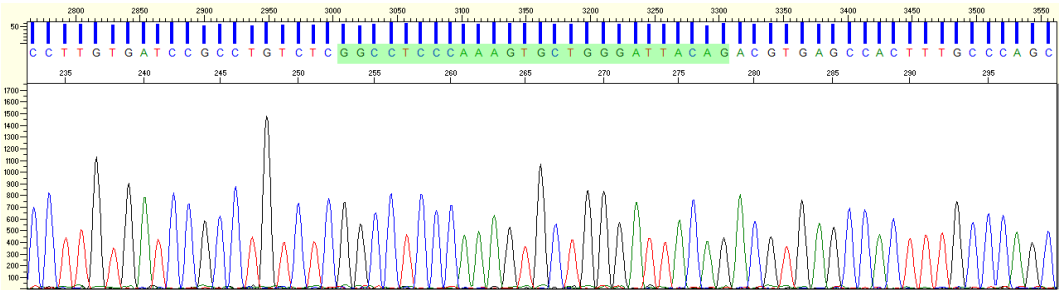


Figure S3. CNV breakpoint sequences. (Top) For each CNV, the "upstream" (telomeric, above) and "downstream" (centromeric, below) reference genomic sequences are aligned with sequence obtained from the breakpoint PCR product (middle). The direction in which the breakpoint PCR product was sequenced (forward vs. reverse) is noted. Blue, bold text indicates identity between the breakpoint sequence and one part of the reference genome; pink, bold text indicates "microhomology" (a short stretch of identity between the upstream and downstream genomic sequences flanking a CNV breakpoint). Known SNPs that could potentially alter interpretation are shown. Some breakpoints are complex and/or can be interpreted in multiple ways (see figure). **(Bottom)** Chromatogram of breakpoint PCR product sequence, with microhomology (if any) highlighted.

A2

Reverse complement sequence

Rendered as an 11 bp insertion

```
chr2:32374397 ACCTCAGGCGATCGCCAGCCTCGGCCTCCCTAAGTGCAAGGATTACAGGCGTGAGCCACCATGCCTGGCCAA 32374468
              ||||||||||||||||||||||||||||| ||||      |       ||  |     | |           ||
A2         501 ACCTCAGGCGATCGCCAGCCTCGGCCTCCC-AAGTAGCTTGGGAGATGGAATCTCTCTCACCCAGGCTTCAT 431
               || | |         ||  |   |   -   |   ||  ||||||||||||||||||||||||||||||||
chr2:32388791 CAGGCATGAGCCACTGCACCCGGCTTTTTT-TTTTTTTTTTTGAGATGGAATCTCTCTCACCCAGGCTTCAT 32388861
```

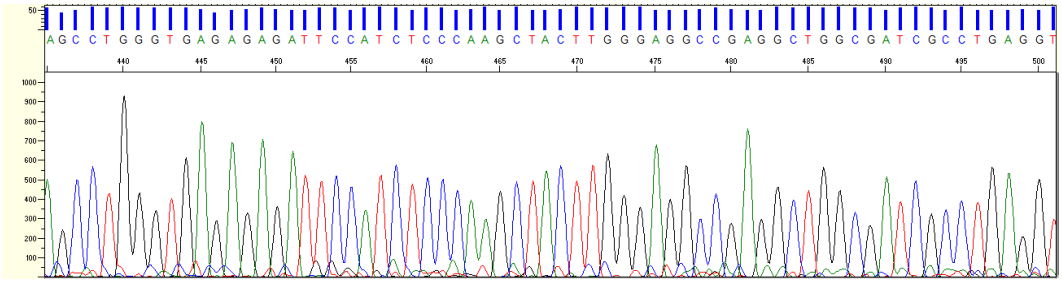
No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

Rendered as template switching x 2 (with a locus ~70 kb away)

```
A2         501 ACCTCAGGCGATCGCCAGCCTCGGCCTCCC-AAGTAGCTTGGGAGATGGAATCTCTCTCACCCAGGCTTCAT 431
              ||||||||||||||||||||||||||||||| ||||      |       ||  |     | |           ||
chr2:32374397 ACCTCAGGCGATCGCCAGCCTCGGCCTCCCTAAGTGCAAGGATTACAGGCGTGAGCCACCATGCCTGGCCAA 32374468
               ||    |  |  ||||| |||||||||||||||||||          | |  ||||  |  |   |
chr2:32303336 GTTCAAGTGATTGTCGTGCCTCAGCCTCCC-AAGTAGCTTGGGATTACAGGCATGTGCCACCATGCCTGGCT 32303406
               || | |         ||  |   |   -   |   ||  ||||||||||||||||||||||||||||||
chr2:32388791 CAGGCATGAGCCACTGCACCCGGCTTTTTT-TTTTTTTTTTTGAGATGGAATCTCTCTCACCCAGGCTTCAT 32388861
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

This potential FoSTeS x 2 mechanism was rendered by searching for the 11bp insertion +/- 1 bp on either side using the UCSC "Short Match" within 200 kb of the CNV, and the middle genomic sequence was the only result.



A3

Reverse complement sequence

```
chr2:32378599 GGTCTCGAACTGCTGACCTCAGGTGATCCACCCGCCTTGGCCTCCCAAAGTGCGGGGATTACAGGTG 32378665
              |||||||||||||||||||||||||||||||||||||| ||| ||||||||||| ||||||||||  |
A3         342 GGTCTCGAACTGCTGACCTCAGGTGATCCACCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAGACG 408
              |||||| | || |||||||   ||||||| ||||||||||||||||||||||||||||||||||||
chr2:32403232 GGTCTCAATCTCCTGACCTT--GTGATCCGCCCGCCTCGGCTTCCCAAAGTGCTGGGATTACAGACG 32403296
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

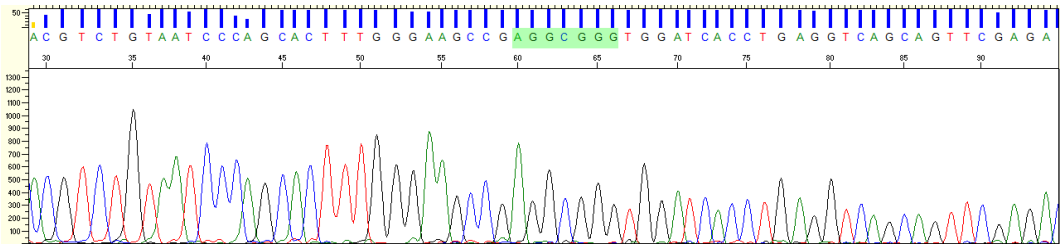


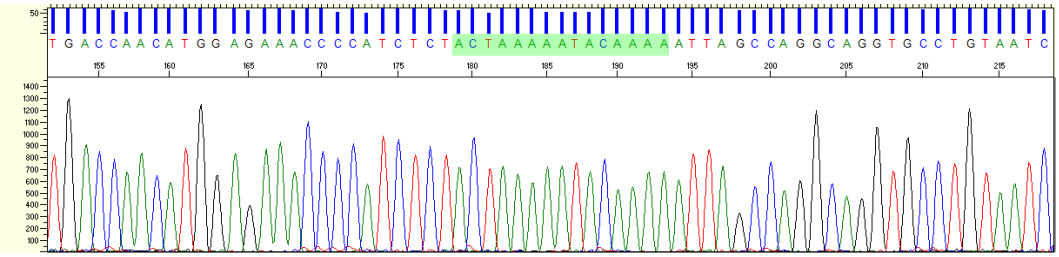
Figure S3, cont'd.

A4

```
Reverse complement sequence

Chr2:32378537 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTGGAGACGGGGTTTCACCATGTT 32378589
              ||-----|||||||||||||||||||||||||||| |||| |||||||| |||||||
A4         209 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 157
              ||     | ||  ||  ||||| ||||||||||||||||||||||||||||||||||||
Chr2:32381252 CATGCGCCACTACCCCAGCTAA-TTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 32381310

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).
```



A5

```
Reverse complement sequence

chr2:32378644 CAAAGTGCGGGGATTACAGGTGTGAGCCACCGTGCTTGGCCAAAATTTTTATTAATTTTT 32378703
              |||||||||||||||||||||||||||||||||| ||||| |  ||||     ||||||
A5         299 CAAAGTGCTGGGATTACAGGCGTGAACCACCGTGCCTGGCCCATGTTTTCTC--ATTTTT 242
              |||||||| |||||||||| |||| ||||||||||||||||||||||||--||||||
chr2:32397022 CAAAGTGCGGGGATTACAGGTGTGAGCCACCGTGCCTGGCCCATGTTTTCTC--ATTTTT 32397079

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).
```

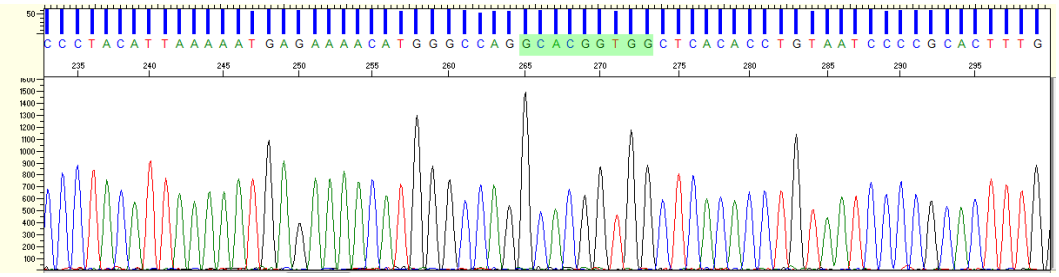


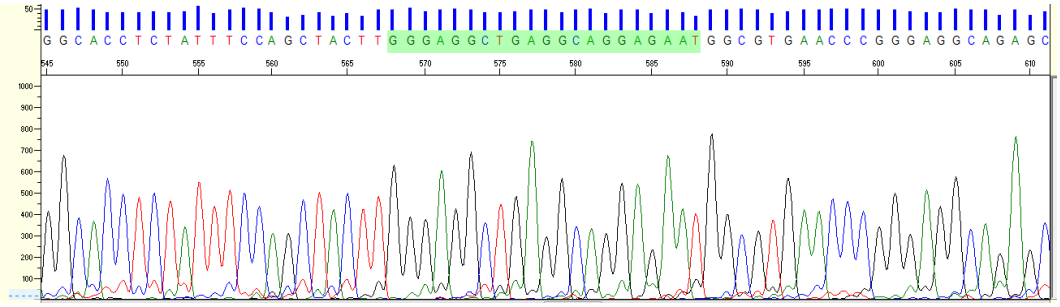
Figure S3, cont'd.

A6

Reverse complement sequence

```
Chr2:32372785 CTGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCTAGTAGCTGGGACTACAGG 32372844
              ||||||||||||||||||||||||||||||||||||||| |||||||| | || |||
A6         608 CTGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGAAATAGAGG 549
              |||| |||| |||||| || ||||||||||||||||||||||||||||||||||||||
Chr2:32398527 CTGCTTCCCTGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGAAATAGAGG 32398586
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).



A8

Reverse complement sequence

```
Chr2:32378537 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTGGAGACGGGGTTTCACCATGTT 32378589
              ||-----|||||||||||||||||||||||||||||| |||| |||||||| |||||||
A8         511 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 459
              ||      | ||  ||  ||||| ||||||||||||||||||||||||||||||||||||
Chr2:32381252 CATGCGCCACTACCCCAGCTAA-TTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 32381310
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

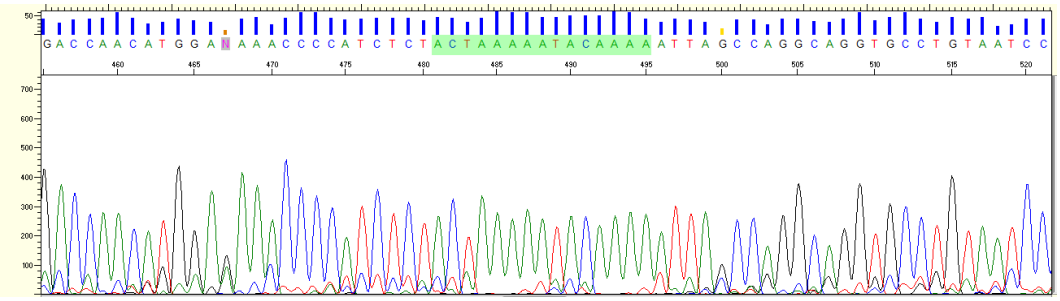


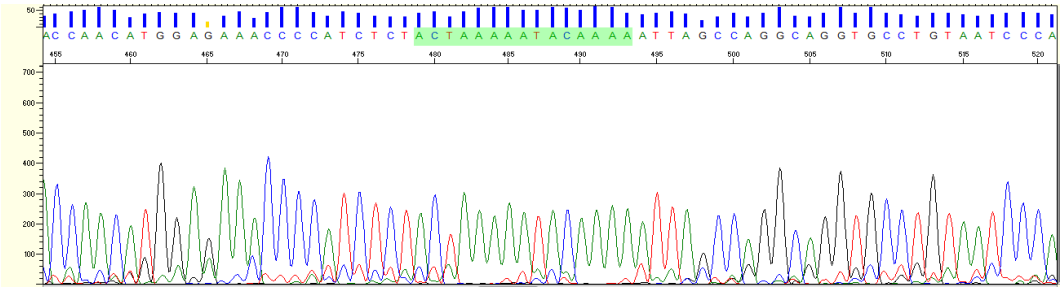
Figure S3, cont'd.

A9

Reverse complement sequence

```
Chr2:32378537 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTGGAGACGGGGTTTCACCATGTT 32378589
              ||-----|||||||||||||||||||||||||||| |||| |||||||| |||||||
A9         509 CA-----C-CT-GCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 457
              ||     | ||  ||  ||||| |||||||||||||||||||||||||||||||||||
Chr2:32381252 CATGCGCCACTACCCCAGCTAA-TTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTT 32381310
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).



A10

Forward sequence

```
chr2:32371786 TAGCCAGGCGTGGTTGTGCGCACCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGTATCGCTTGAA 32371855
              ||||||||||||||||||||||||||||||||||||||||||||||||||          ||||||||||
A10        441 TAGCCAGGCGTGGTTGTGCGCACCTGTAATCCCAGCTACTCAGGAGGCTGATA-------ATCGCTTGAA 503
              ||||  || |||||    | ||||||||||||||||||||||||||||||||-------||||||||||
chr2:32365545 TAGCTGGGTGTGGTCACACACACCTGTAATCCCAGCTACTCAGGAGGCTGATA-------ATCGCTTGAA 32365607
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

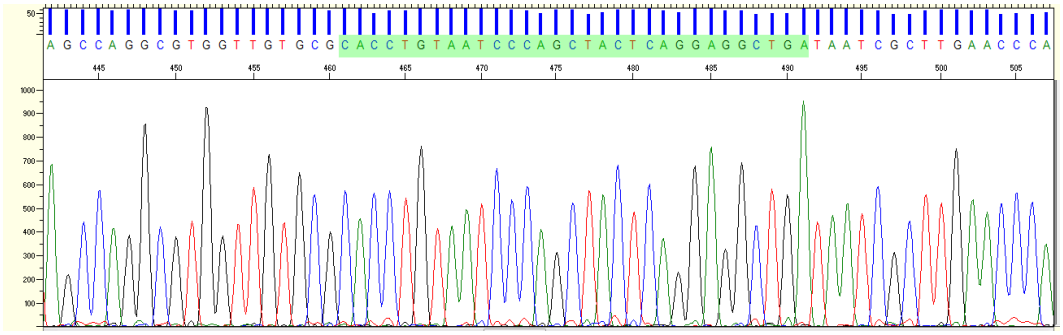


Figure S3, cont'd.

BAB 3200

Forward sequence

Rendered as a 16 bp insertion

```
Chr2:32260417 AATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGATCAAGACCATCCTGGCCA 32260484
              |||||||||||||||||||||||| |||  ||||||  |  | | |  |   |    |   || ||
BAB 3200  173 AATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGCTGAAATTACTGTTTTTATAGGTCA 240
                 |  ||  ||  |       |     | ||   |||||| ||||||||||||||||||||||||
Chr2:32347647 TCCCAAAGTGCTGGGATTACAGGTATGAGCCACCACACCTGGGCTGAAATTACTGTTTTTATAGGTCA 32347714
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

Rendered as a template switch x 2 (with a locus ~13 kb away)

```
BAB 3200  173 AATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCACCTGAGCTGAAATTACTGTTTTTATAGGTCA 240
              ||||||||||||||||||||||||| |||  ||||||  |  | | |  |   |    |   || ||
Chr2:32260417 AATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGATCAAGACCATCCTGGCCA 32260484
              ||||| |||||||||||||||| |||||||||||||||||||| | |       |     |
Chr2:32247023 AATCCTAGCACTTTGGGAGGCTGAGGTGGGCAGATCACCTGAGGTCAGGAATTCGAGACCAGCCTGGC 32247090
                 |  ||  ||  |       |     | ||   |||||| ||||||||||||||||||||||||
Chr2:32347647 TCCCAAAGTGCTGGGATTACAGGTATGAGCCACCACACCTGGGCTGAAATTACTGTTTTTATAGGTCA 32347714
```

No known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).

This potential FoSTeS x 2 mechanism was rendered by searching for the 16bp insertion +/- 1 bp on either side using the UCSC "Short Match" feature to look within 200 kb of the CNV, and the middle genomic sequence was the most closely located result.

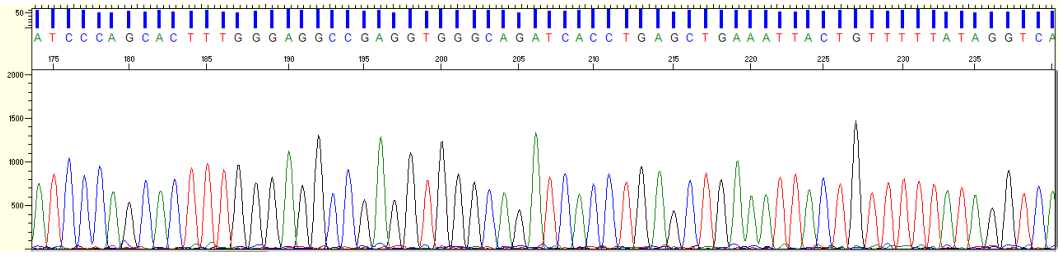


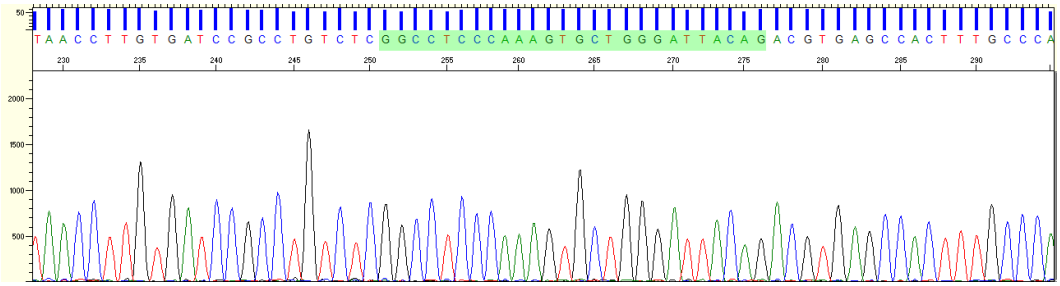
Figure S3, cont'd.

BAB 3327

Reverse complement sequence

```
Chr2:32378859 CTTTAGGCTGGGCAAAGTGGCTCACGTCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTTAGG 32378937
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||   ||||||||||   ||||||||
BAB 3327  303 CTTTAGGCTGGGCAAAGTGGCTCACGTCTGTAATCCCAGCACTTTGGGAGGCCGAGACAGGCGGATCACA--AGGTTAGG 226
                | ||||||||   |||||||||| ||||||||||||||||||||||||||||||||||| |||||||--|||| |||
chr2:32384360 TAATTGGCTGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGACAGGCAGATCACA--AGGTCAGG 32384437
                rs187211683* * rs4519572                                rs4407291*             *rs4530415
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP135 via UCSC Genome Browser).



BAB 5112

Forward sequence

```
chr2:32327123 ATTACCTTGGGCAGTGTGGCCATTTTCATGATATTGATTCTTCCTATCCATAAGCATGGAATGTTCTTCCATTTGTCTGTG 32327203
              ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||| |||||||||| ||||
BAB 5112  377 ATTACCTTGGGCAGTGTGGCCATTTTCATGATATTGATTCTTCCTATCCATGAGCATGGAATGTTTTTCCATTTGTTTGTG 457
              ||||| ||||||| | ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32886415 ATTACTTTGGGCAATATGGCCATTTTCACCATATTGATTCTTCCTATCCATGAGCATGGAATGTTTTTCCATTTGTTTGTG 32886495
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

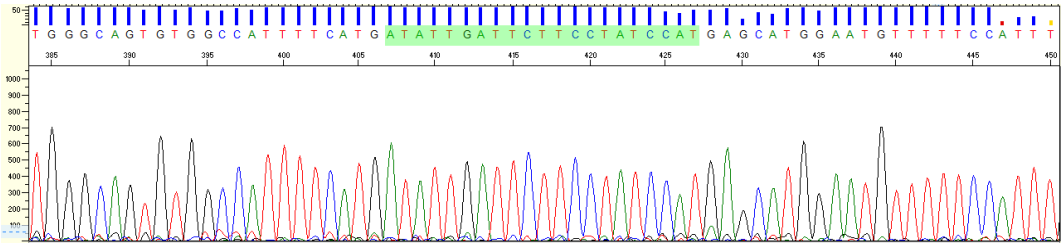


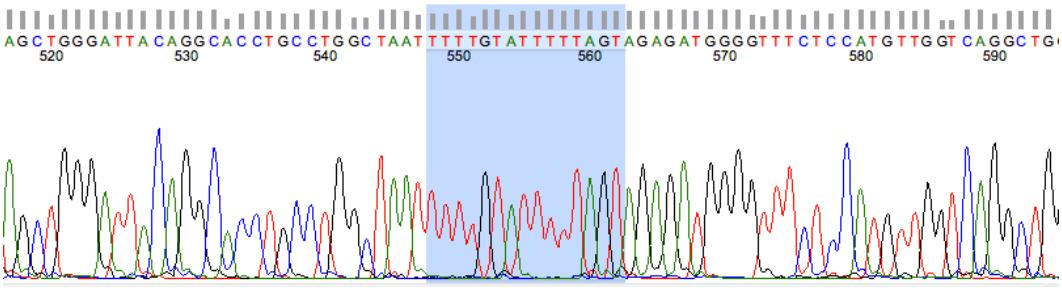
Figure S3, cont'd.

SPAST 1

Forward sequence

```
Chr2:32378515 AGCTGGGATTACAGGCACCTGCCTG-------GCTAATTTTTGTATTTTTAGTGGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCG 32378605
              |||||||||||||||||||||||-------|||||||||||||||||||| |||| |||||||| ||||||||| |||||||||| |
KCL133376 510 AGCTGGGATTACAGGCACCTGCCTG-------GCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCTTG 600
              |||||||||||||||||   |||         ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Chr2:32381237 AGCTGGGATTACAGGCATGCGCCACTACCCCAGCTAA-TTTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCTTG 32381326
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



SPAST 3

Forward sequence

```
Chr2:32726526 TCCATTCGGGTTATGTGTAGAGGGTTTAGACAGCTATTATATATTCATTCTTCAGCAGTTTTGGAAAAAGTG 32726597
              |||||||||||||||||||||||||||||||||||| |       | |       |||   ||
KCL186204 938 TCCATTCGGGTTATGTGTAGAGGGTTTAGACAGCTATTTTTGTACTTTAGTCTTAATCCAGTGGTTGAAAGA 1009
              |   |       || | |                 ||||||||||||||||||||||||||||||||||
Chr2:32294329 TATCTATCTTAAATTTTTTCTTTAGGAGATTGAATATTTTTGTACTTTAGTCTTAATCCAGTGGTTGAAAGA 32294400
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

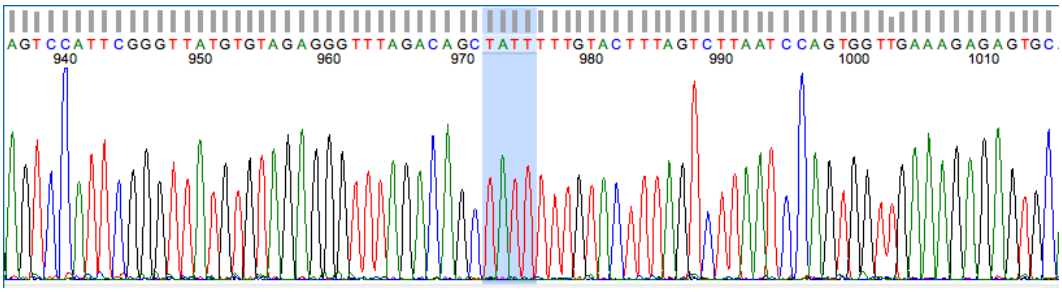


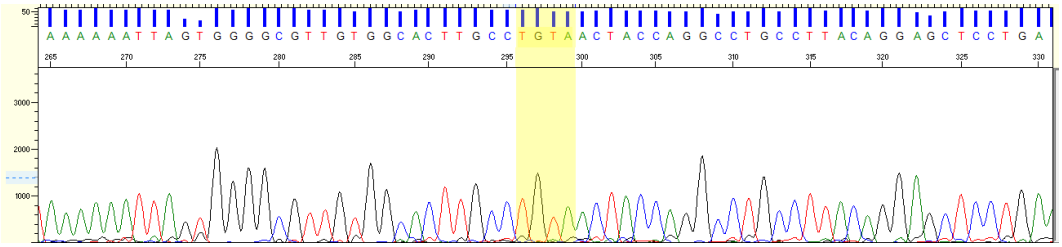
Figure S3, cont'd.

Spain 1

Reverse complement sequence

```
chr2:32329691 CAGGAGCTCCTGTAAGGCAGGCCTGGTAGTTACAAAATCTGTCAGCATTTGCTTGTCTTTAAAG 32329754
              ||||||||||||||||||||||||||||||||||||          |  ||    |       |
Spain 1   266 CAGGAGCTCCTGTAAGGCAGGCCTGGTAGTTACAGGCAAGTGCCACAACGCCCCACTAATTTTT 329
                    |||          ||     |    ||||||||||||||||||||||||||||||||
chr2:32373110 TCTTGTCTCAAACTTTTTAGTAGCTGGGACTACAGGCAAGTGCCACAACGCCCCACTAATTTTT 32373173
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Spain 2

Forward sequence

```
chr2:32362794 CTTTACTGGATGGGTTACGTGGGTATATTGCACCCAGGTAGTGAGCATAGGGTGCAGTAG Chr:32362853
              ||||||||||||||||||||||||||||||||||||     |       |              ||
Spain 2   568 CTTTACTGGATGGGTTACGTGGGTATATTGCTGGGATTACAGGCATGCGCCACCATGTCC 627
               |         ||  | |         | ||||||||||||||||||||||||||||||
chr2:32817764 GTGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGGATTACAGGCATGCGCCACCATGTCC Chr:32817823
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

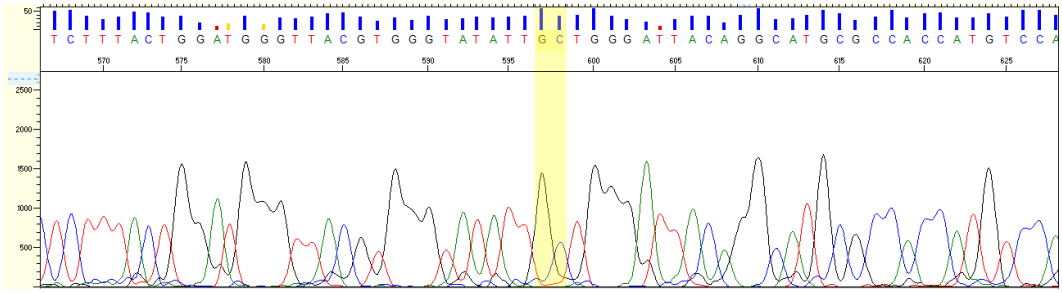


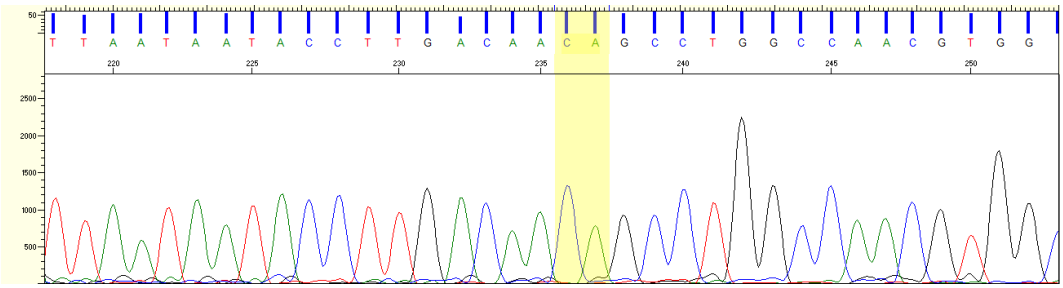
Figure S3, cont'd.

Spain 3

Reverse complement sequence

```
Chr2:32313141 AAGATGGGGTTTTGCCACGTTGGCCAGGCTGGTTTCGAACTCCTGATCTCAAGTGATCTG Chr2:32313200
              ||||||||||||||||||||||||||||||| | || |  |  |  |   |  |    |
Spain 3   266 AAGATGGGGTTTTGCCACGTTGGCCAGGCTGTTGTCAAGGTATTATTAATATTTATGTTA 207
              ||       |  |   |  |    ||   ||||||||||||||||||||||||||||
chr2:32408889 AAATGAACATAATTTTATCTCACTCAAAATGTTGTCAAGGTATTATTAATATTTATGTTA Chr2:32408948
      rs141605765*
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Spain 4

Reverse complement sequence

```
chr2:32342663 TGATGTTGGCTCACTGCAACCTCTGCCTTTGGGGTTCA-AGTGATTCTCATGCCTCATCCTCCTGAGT 32342729
              |||||||||||||||||||||||||||||||||||| | | |||||  ||| |      |    |||
Spain 4   112 TGATGTTGGCTCACTGCAACCTCTGCCTTTGGGGTTCAGATTCATTCTTTTGCATGTGGATAACCAGT 45
                |  |   |   |            || ||||||||||||||||||||||||||||||||||||
chr2:32403316 AAGTTATTTTTTTTTTAATGGTGTGAGATTGGGGGTTCAGATTCATTCTTTTGCATGTGGATAACCAGT 32403383
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

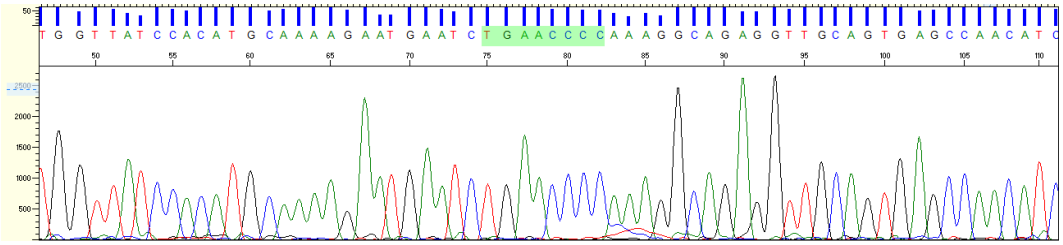


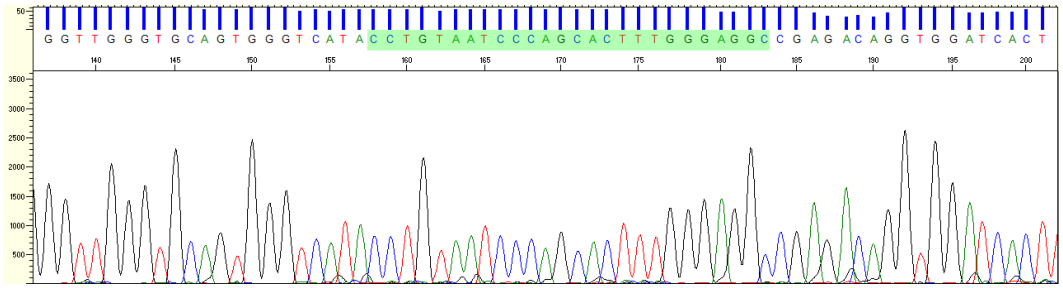
Figure S3, cont'd.

Spain 5

Reverse complement sequence

```
Chr2:32282809 CTCCTGACCTCAAGTGATCCACCTGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTTAGCCACAGCATCCAGCCTTAAGCATG Chr2:32282894
              ||||||||||||||||||||||||||||||||||||||||||||||||||  | | |||| ||| ||| ||    |
Spain 5   213 CTCCTGACCTCAAGTGATCCACCTGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGTATGACCCACTGCACCCAACCCATACTCAA 128
              |||||||||| ||||||||| |||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Chr2:32295360 CTCCTGACCTGAAGTGATCCGCCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGTATGACCCACTGCACCCAACCCATACTCAA Chr2:32295445
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Spain 6

Reverse complement sequence

```
Chr2:32302017 TCACTGCAACCTCTGTTTCCCAGGTTCAAACGATTCTCCTGCCTCAGCCTCCCAAGTAGC Chr2:32302076
              ||||||||||||||||||||||||||||||      |     ||     | || ||  |
Spain 6   204 TCACTGCAACCTCTGTTTCCCAGGTTCAAAGCTGAGTATCTTCTACAATTTCCCAGATGA 145
                | |  |        |       ||  |  ||||||||||||||||||||||||||||||
chr2:32472132 CAAGTAAACAACAGTTAGGGTGGGAACGGAGCTGAGTATCTTCTACAATTTCCCAGATGA Chr2:32472191
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

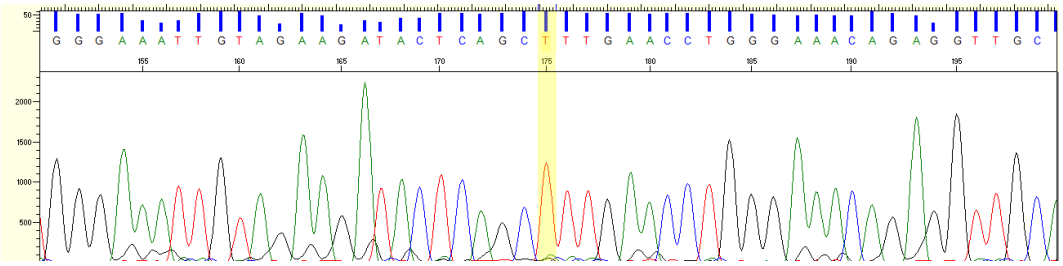


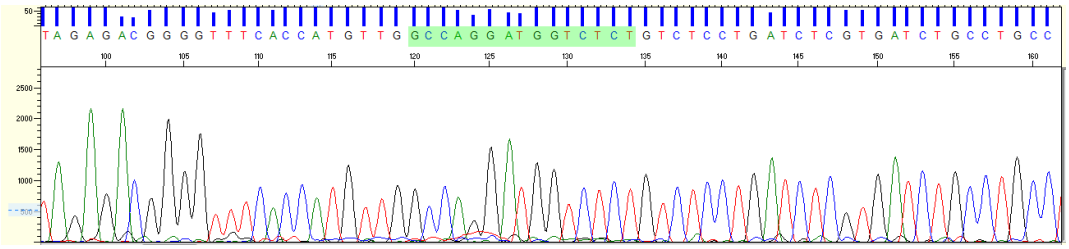
Figure S3, cont'd.

Spain 7

Forward sequence

```
chr2:32356900 TTTTGTA-TTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCCTGACCTTGTGATCCGCCCGCATCTACCTCCCA 32356989
              |||||||-|||||||||||||||||||||||||||||||||||||||||||| |||||||| || |||||| ||| || ||  |||||||
Spain 7    83 TTTTGTA-TTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTCTGTCTCCTGATCTCGTGATCTGCCTGCCTCAGCCTCCCA 172
              ||||||| ||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32372870 TTTTGTACTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCTGTCTCCTGATCTCGTGATCTGCCTGCCTCAGCCTCCCA 32372960
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Spain 8

Reverse complement

```
chr2:32302096 CACCATGCCTGGCTAA--TTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATCTTGGCCAGGCTGATCTCGAACTCTTGACCTCATGA 32302182
              |||||||||||||||--|||||||||||||||||||||||||| | || |||||| || |||||| ||| ||||||| |||
Spain 8   496 CACCATGCCTGGCTAA--TTTTTGTATTTTTAGTAGAGATGGGGTTTCATCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGA 410
              ||||| ||| ||||||  |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
chr2:32351015 CACCACGCCCGGCTAAttttttttGTATTTTTAGTAGAGACGGGGTTTCATCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGA 32351103
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

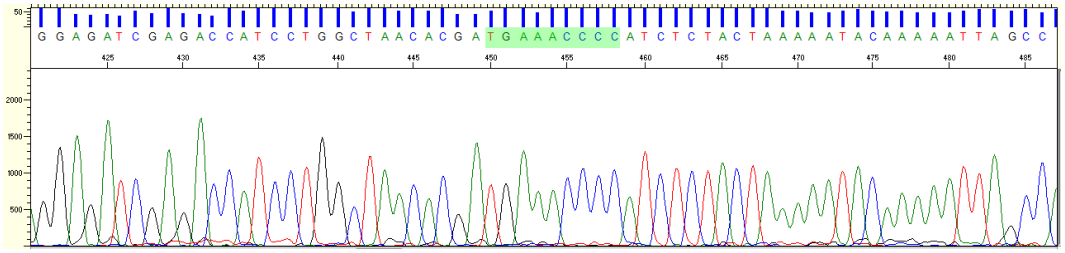


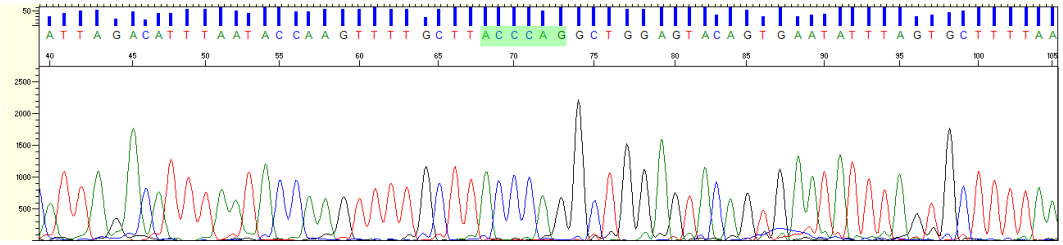
Figure S3, cont'd.

Spain 9

Forward sequence

```
                                                          rs181046857*
chr2:32340370 TCATTAGACATTTAATACCAAGTTTTGCTTACCCAGTCTGTTTTATATAATGCTGTATATTTCATT 32340435
              ||||||||||||||||||||||||||||||||||| ||| || | |  | |  ||
Spain 9    38 TCATTAGACATTTAATACCAAGTTTTGCTTACCCAGGCTGGAGTACAGTGAATATTTAGTGCTTTT 103
               | |       |     || | |||| || ||||||||||||||||||||||||||||||||||
chr2:32345390 CCTTCTACTCCTGTCCTCCCACTTTTTCTCACCCAGGCTGGAGTACAGTGAATATTTAGTGCTTTT 32345455
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



JKF B

Forward sequence

```
chr2:32371871 TGCAGTAAGCCGAGCACTGCAGCCTGGGCAACAGAGCAAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAA 32371944
              ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
JKF B     226 TGCAGTAAGCCGAGCACTGCAGCCTGGGCAACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAAANNNNNNNNN 299
               |       |   |||||  |||||||||| ||||||||||||||||||||||||||||||
chr2:32373830 AGATCGTGCCACTGCACTCCAGCCTGGGCGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAATAGCTGGGC 32373903
                  *rs9677320                                                *Multiple SNPs
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

The low-complexity region 3' of the breakpoint (poly-A) renders the sequence flanking the breakpoint somewhat less certain than in other subjects. Thus, this breakpoint may be considered provisional. This position of the breakpoint coincides with the length of the breakpoint PCR product (Supplemental Fig. S4).

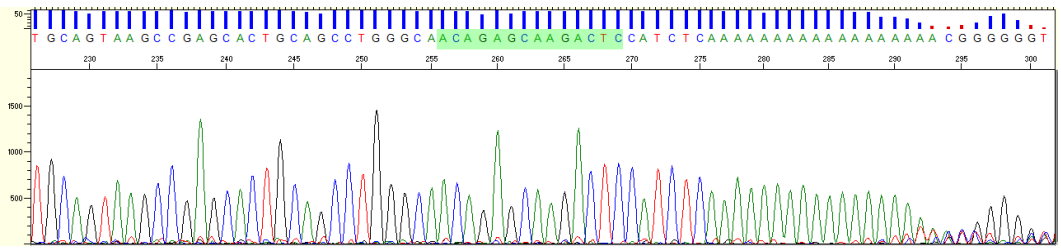


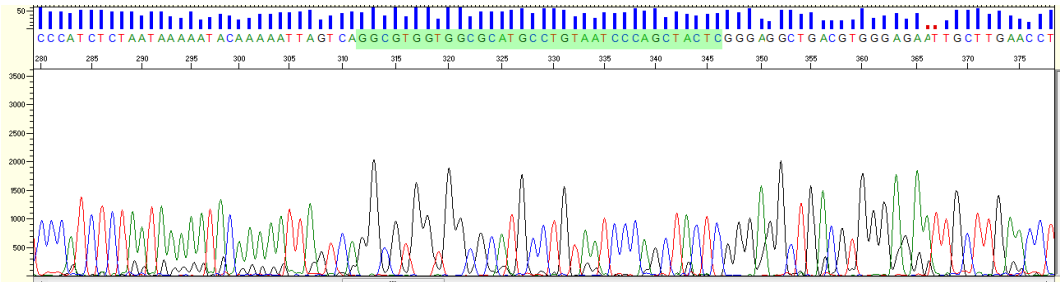
Figure S3, cont'd.

JKF C

Reverse complement sequence

```
                                                                                *rs188067254
Chr2:32359860 GTTCAAGCAATTCTCCCACGTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTAGTTTTAGTAGACACA Chr2:32359954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||| |||| |
JKF C     376 GTTCAAGCAATTCTCCCACGTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCACGCCTGACTAATTTTTGTATTTTTATTAGAGATG 282
              |||||||||||||| |||||||||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Chr2:32353159 GTTCAAGCAATTCTGCCACGTCAGCTTCCTGAGTAGCTGGGATTACAGGCATGCGCCACCACGCCTGACTAATTTTTGTATTTTTATTAGAGATG Chr2:32353253
                                     *rs216540
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 3

Forward sequence

```
chr2:32291741 aaaaaaaGTAGGTCATTTTTGGCTGGGCACGGTGGCTCATGCCTGTAATTCCAGCACTTTGGGAGGCTGAGGTGGGTGG chr2:32291819
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||  |||  |||||
Ruhr 3    369 AAAAAAAGTAGGTCATTTTTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGTGG 447
                  |     || |   | ||   |||| ||    ||||||||||||||||||||||||||||||||||||||||||
chr2:32357157 TTTACTTTTAAGAAAGTTGAGGCTAGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGTGG chr2:32357235
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

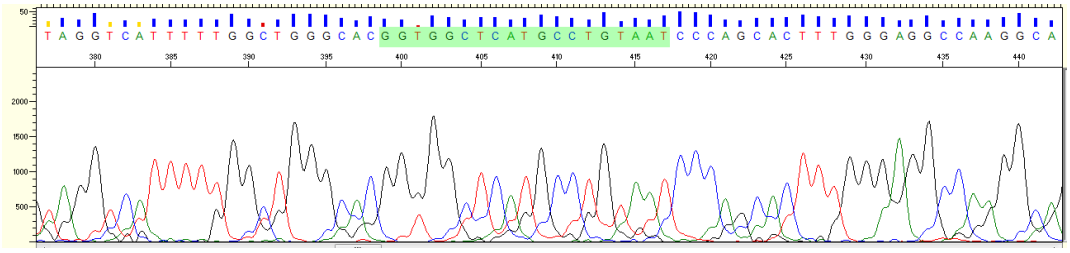


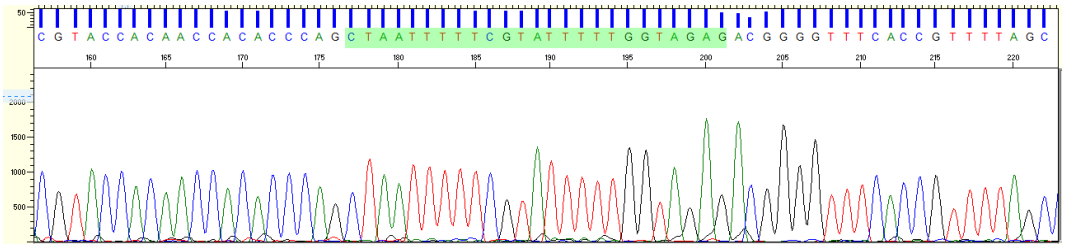
Figure S3, cont'd.

Ruhr 4

Forward sequence

```
chr2:32282718 ACCACAGACACGTACCACAACCACACCCAGCTAATTTTTCGTATTTTTGGTAGAGGTGGGGTTTCACCATGTTGGCCAGGCTGGT 32282802
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||||| | || ||| || ||||
Ruhr 4    147 ACCACAGACACGTACCACAACCACACCCAGCTAATTTTTCGTATTTTTGGTAGAGACGGGGGTTTCACCGTTTTAGCCGGGATGGT 231
              || || ||| |||||||||| ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32303212 ACTACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGACCTCGT 32303296
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 5

Forward sequence

```
chr2:32291741 aaaaaaaGTAGGTCATTTTTGGCTGGGCACGGTGGCTCATGCCTGTAATTCCAGCACTTTGGGAGGCTGAGGTGGGTGG chr2:32291819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||  |||  |||||
Ruhr 5    368 AAAAAAAGTAGGTCATTTTTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGTGG 446
                  |    || |  | ||  |||| ||   ||||||||||||||||||||||||||||||||||||||||||||||
chr2:32357157 TTTACTTTTAAGAAAGTTGAGGCTAGGTGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGTGG chr2:32357235
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

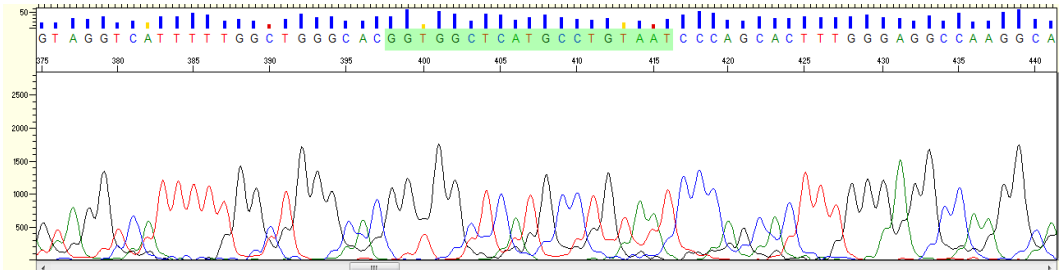


Figure S3, cont'd.

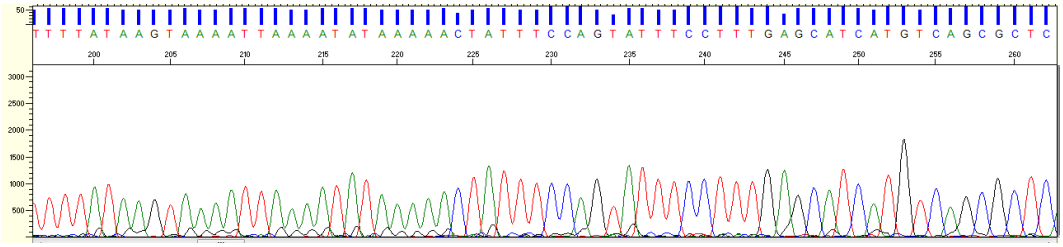
Ruhr 6

Reverse complement sequence

```
chr2:32332644 GCTGACATGATGCTCAAAGGAAATACTGGAGTATTTTGCATTTTGGATTTTTGGGTTAAGG Chr2:32332704
              ||||||||||||||||||||||||||||||      ||     | |      |||    |||
Ruhr 6    258 GCTGACATGATGCTCAAAGGAAATACTGGAAATAGTTTTTATATTTTAATTTTACTTATAA 198
                | |   |     |     |      |        ||||||||||||||||||||||||||||
chr2:32340904 AGTAAGTTTTGCCATCTAAATGttttattttatagtttttatattttaattttaCTTATAA Chr2:32340964
              *rs201434161  *rs189951674
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

Low complexity marked with lower case letters.



Ruhr 7

Reverse complement sequence

```
chr2:32376073 GTGCAATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCGCCACTATGC 32376172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Ruhr 7    599 GTGCAATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACTA-CC 501
              | |||||||||| |||||  |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||-||
chr2:32381167 GCGCAATCTCGACTCACCACAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACTA-CC 32381265
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

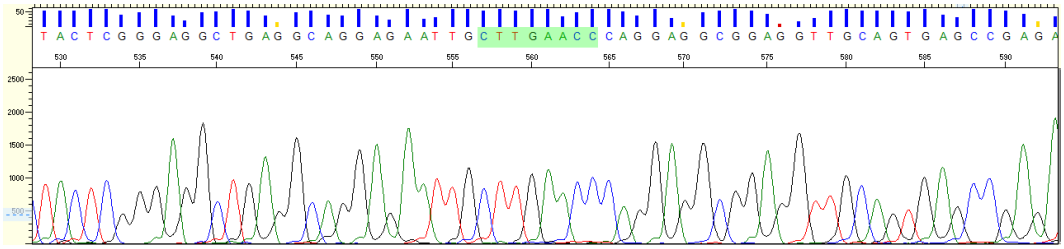


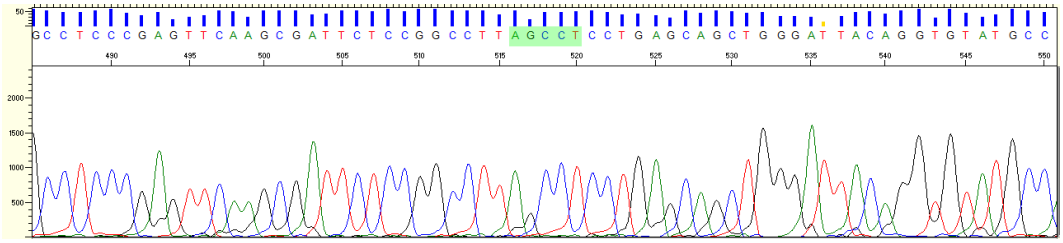
Figure S3, cont'd.

Ruhr 9

Forward sequence

```
Chr2:32288184 CCTCCCGAGTTCAAGCGATTCTCCGGCCTTAGCCTTCCAACCAGCTGGGACTACAGGCGCGCGCC 32288248
              |||||||||||||||||||||||||||||||||||| |  | ||||||||| |||||| |   |||
Ruhr 9    486 CCTCCCGAGTTCAAGCGATTCTCCGGCCTTAGCCTCCTGAGCAGCTGGGATTACAGGTGTATGCC 550
              |||||||  ||||||||||||||| |||| ||||||||||||||||||||||||||||||||||
Chr2:32292218 CCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGCAGCTGGGATTACAGGTGTATGCC 32292282
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 10

Forward sequence

```
chr2:32288184 CCTCCCGAGTTCAAGCGATTCTCCGGCCTTAGCCTTCCAACCAGCTGGGACTACAGGCGCGCGCC 32288248
              |||||||||||||||||||||||||||||||||||| |  | ||||||||| |||||| |   |||
Ruhr 10   487 CCTCCCGAGTTCAAGCGATTCTCCGGCCTTAGCCTCCTGAGCAGCTGGGATTACAGGTGTATGCC 551
              |||||||  ||||||||||||||| |||| ||||||||||||||||||||||||||||||||||
chr2:32292218 CCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGCAGCTGGGATTACAGGTGTATGCC 32292282
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

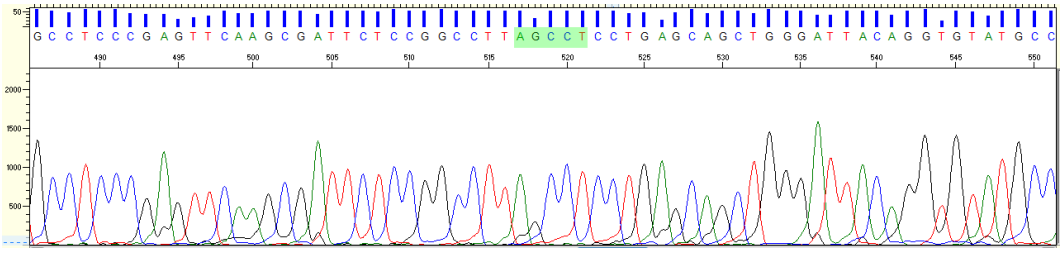


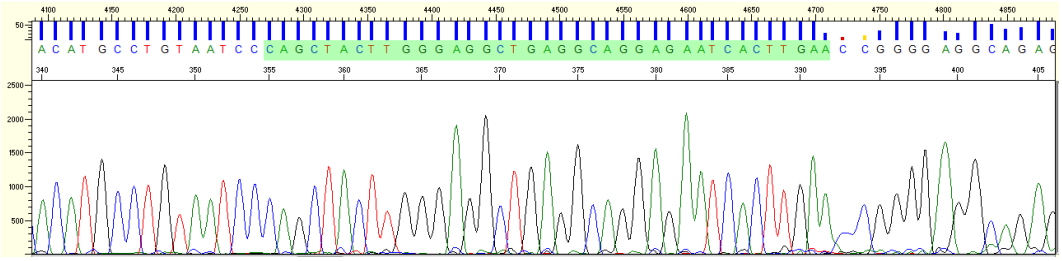
Figure S3, cont'd.

Ruhr 11

Reverse complement sequence

```
                                                                                                rs148709324*
chr2:32308107 CACAGCTCACTGCAACCTCTGCCTCCCCG-GTTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGTGATTACAGGTGCCCGCCACCACAACCAGC 32308204
              ||||||||||||||||||||||||||||-||||||||||||||||||||||||||||||||||||| |||||||||     |||||  |    |||||
Ruhr 11   422 CACAGCTCACTGCAACCTCTGCCTCCCCG-GTTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATGTGCCACTATGCCCAGC 325
              |  |||||||||||||||||| |||| |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32341601 CTTAGCTCACTGCAACCTCTGCGTCCCAGAATTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATGTGCCACTATGCCCAGC 32341699
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 12

Reverse complement sequence

```
chr2:32274877 TTCTGTATTTTTAGTGGAGACGGAGTTTCTCCATGTTGGTCAGGCTGGTCAGAAACTCCCGACCTCAGG 32274945
              |||||||||||||||||||||||||||||||||||||||||| ||| ||||||| ||||||  |||||||||
Ruhr 12   205 TTCTGTATTTTTAGTGGAGACGGAGTTTCTCCATGTTGGCCAGACTGGTCACGAACTCCTGACCTCAGG 137
              || |||||||||||| |||| || ||||| ||||||||||||||||||||||||||||||||||||||||
chr2:32291212 TTTTGTATTTTTAGTAGAGATGGTGTTTCACCATGTTGGCCAGACTGGTCACGAACTCCTGACCTCAGG 32291280
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

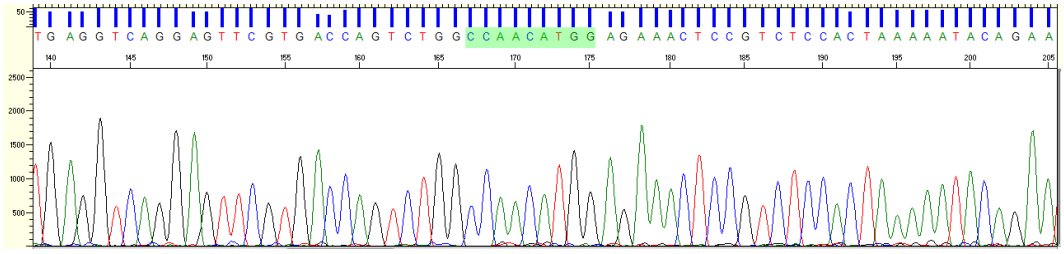


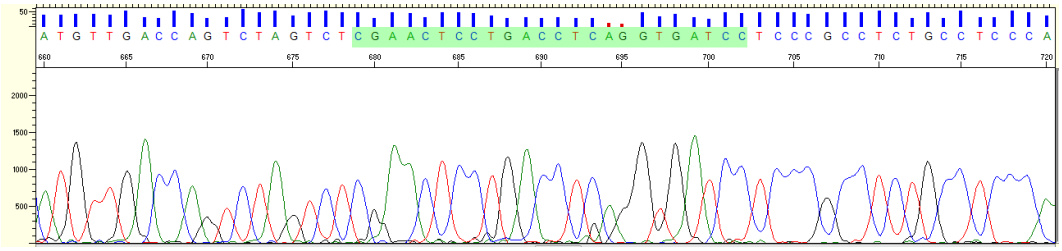
Figure S3, cont'd.

Ruhr 13

Forward sequence

```
chr2:32347580 GGGGTTTTGCCATGTTGACCAGTCTAGTCTCGAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTATG 32347673
              ||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||  |||||||||||||||||||||||||  ||
Ruhr 13   649 GGGGTTTTGCCATGTTGACCAGTCTAGTCTCGAACTCCTGACCTCAGGTGATCCTCCCGCCTCTGCCTCCCAAAGTGCTGGGATTACAGGCGTG 742
              |||||||  ||||||||  ||| ||  || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32392333 GGGGTTTCTCCATGTTGGTCAGGCTGATCGCGAACTCCTGACCTCAGGTGATCCTCCCGCCTCTGCCTCCCAAAGTGCTGGGATTACAGGCGTG 32392426
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 14

Reverse complement sequence

```
chr2:32370384 TTGAACTCCCGGGCTCAAGTGAGCAGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCATGCC 32370460
              ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||  |||||||||| |||
Ruhr 14   275 TTGAACTCCCGGGCTCAAGTGAGCAGCCTGCCTTGGCCTCCCAAAGTGTTGGGATTACAGGTGTGAGCCACCACGCC 199
              | ||||||| |  ||||||||| | |||  ||||||||||||||||||||||||||||||||||||||||||||
chr2:32402060 TCGAACTCCTGACCTCAAGTGATCTGCCCACCTTGGCCTCCCAAAGTGTTGGGATTACAGGTGTGAGCCACCACGCC 32402136
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

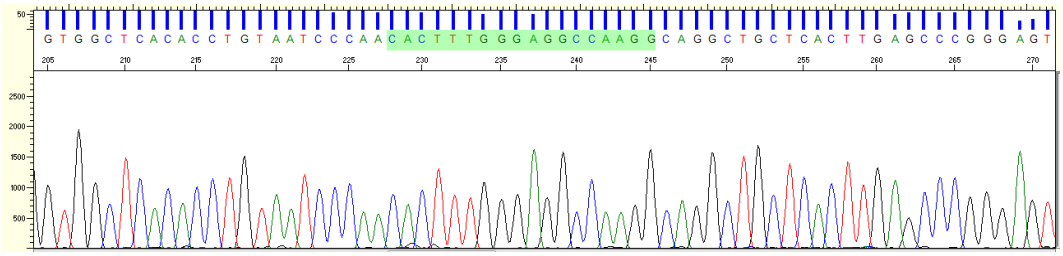


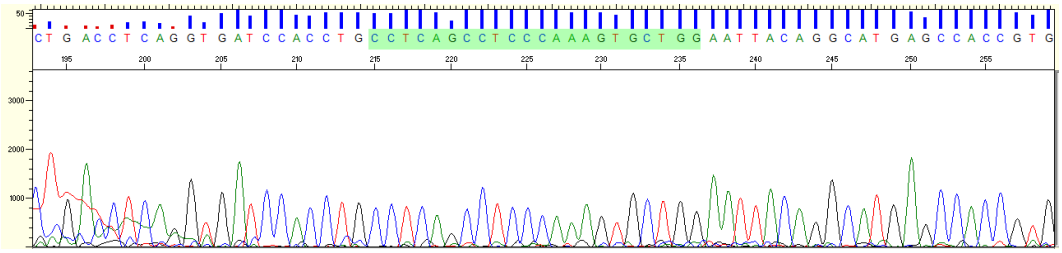
Figure S3, cont'd.

Ruhr 15

Reverse complement sequence

```
chr2:32291761 GGCTGGGCACGGTGGCTCATGCCTGTAATTCCAGCACTTTGGGAGGCTGAGGTGGGTGGATTGCTTGAGCCCAGGAGTTTGA 32291842
              ||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||  | ||||  |||||||| ||
Ruhr 15   266 GGCTGGGCACGGTGGCTCATGCCTGTAATTCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGA 185
              |||||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32364615 GGCTGGGCACAGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTCAGGAGTTCGA 32364696
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 16

Reverse complement sequence

```
chr2:32338390 GTGATCTGCCCGCCTCAGCCTCCCAAAATGCTGGGATTACAGGCGTGAGCTACTGCACCCGGCCTAAAATTGATTAGTTT 32338469
              ||||||||||||||||||||||||||||||||||||||||||||||||||| || |  |||    | || ||  ||
Ruhr 16   241 GTGATCTGCCCGCCTCAGCCTCCCAAAATGCTGGGATTACAGGCGTGAGCCACCGTGCCCAAAGTTAATTTTTTTTTAAA 162
              |||||| ||| |||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
chr2:32340253 GTGATCCGCCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACCGTGCCCAAAGTTAATttttttttttAAA 32340332
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

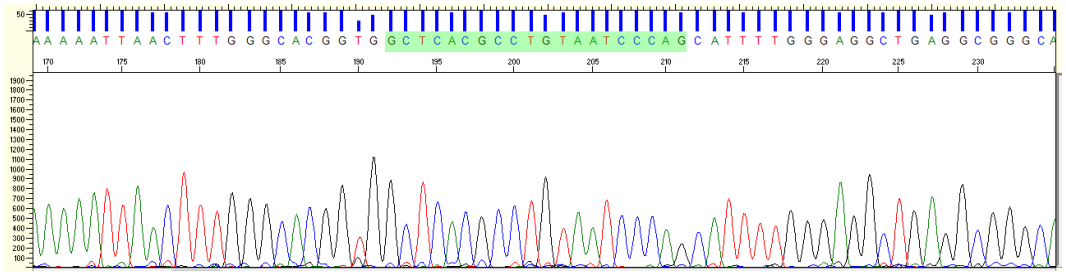


Figure S3, cont'd.

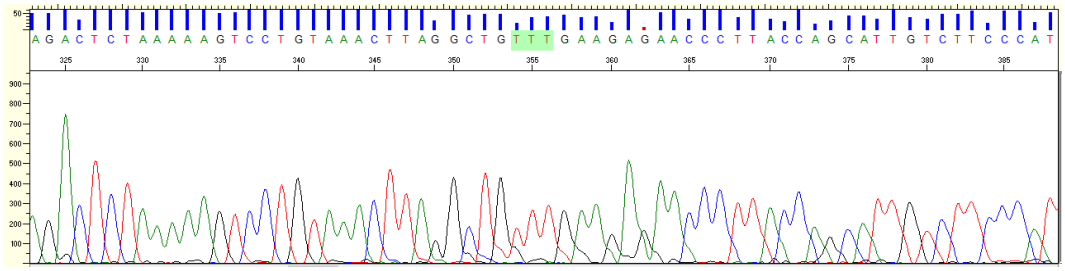
Ruhr 17

Reverse complement

```
chr2:32352077 GGGAAGACAATGCTGGTAAGGGTTCTCTTCAAATTTGAGTTTTCTGTTGAGATATTTGGGATA 32352139
              ||||||||||||||||||||||||||||||||||          | |   |   | ||| |  |
Ruhr 17   386 GGGAAGACAATGCTGGTAAGGGTTCTCTTCAAACAGCCTAAGTTTACAGGACTTTTTAGAGTC 324
                ||     |               ||    ||||||||||||||||||||||||||||||
chr2:32379616 AAGATCTTCAATGAACGTCATCGGCTACAGAAACAGCCTAAGTTTACAGGACTTTTTAGAGTC 32379678
              *rs6730121
```

The 3' end of exon 8 is underlined on the top (telomeric) genomic sequence.

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 18

Forward sequence

```
chr2:32347580 GGGGTTTTGCCATGTTGACCAGTCTAGTCTCGAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGAT 32347663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||  ||||||||||||||||||
Ruhr 18   650 GGGGTTTTGCCATGTTGACCAGTCTAGTCTCGAACTCCTGACCTCAGGTGATCCTCCCGCCTCTGCCTCCCAAAGTGCTGGGAT 733
              |||||||  ||||||||  ||| ||  || ||||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32392333 GGGGTTTCTCCATGTTGGTCAGGCTGATCGCGAACTCCTGACCTCAGGTGATCCTCCCGCCTCTGCCTCCCAAAGTGCTGGGAT 32392416
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

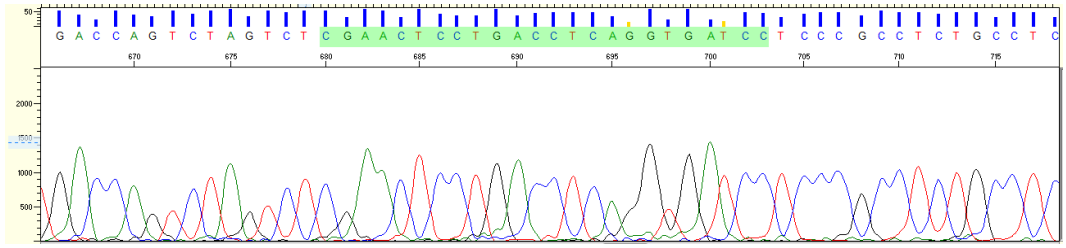


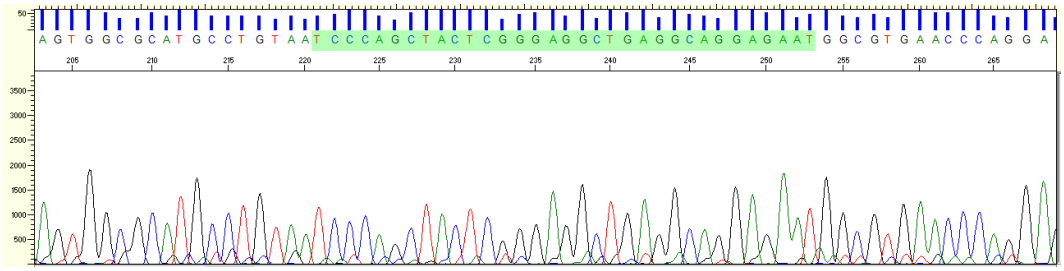
Figure S3, cont'd.

Ruhr 21

Reverse complement sequence

```
chr2:32360273 ACTGCAAGCTCCACCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTGCCACCACGCCTGGCTAAT 32360366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||     ||||| || || | |||||
Ruhr 21   283 ACTGCAAGCTCCACCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACTACCCCAG-CTAAT 191
              ||  ||| |||| |||||  |||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||-|||||
chr2:32381182 ACCACAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCGCCACTACCCCAG-CTAAT 32381274
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

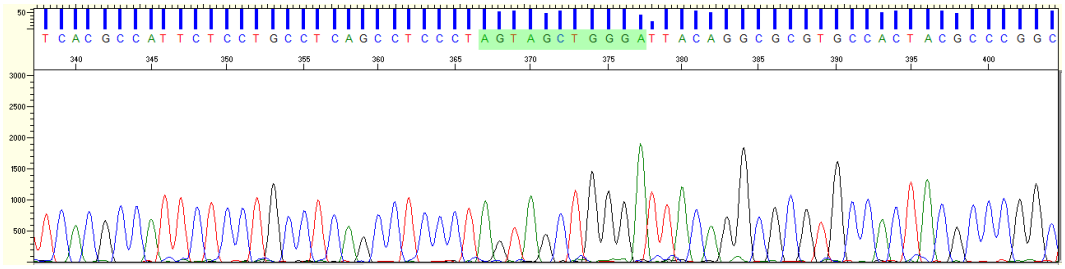


Ruhr 23

Forward sequence

```
chr2:32372797 TTCACGCCATTCTCCTGCCTCAGCCTCCCTAGTAGCTGGGACTACAGGCGCCCACCACCACGCCTGGCTAA 32372867
              |||||||||||||||||||||||||||||||||||||||||| |||||||||    |||| ||||| ||||||
Ruhr 23   337 TTCACGCCATTCTCCTGCCTCAGCCTCCCTAGTAGCTGGGATTACAGGCGCGTGCCACTACGCCCGGCTAA 407
              |||| || |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
chr2:32390371 TTCAAGCGATTCTCCTGCCTCAGCCTCCAGAGTAGCTGGGATTACAGGCGCGTGCCACTACGCCCGGCTAA 32390441
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 26

Sequencing failed to identify the breakpoint sequence.

Figure S3, cont'd.

Ruhr 27

Forward sequence

```
Ruhr 27   223 TGTTTTGTTTGTTTGTTTGTTTGTTTGTTTGTTTGAGACGGAGTCTCGCGC-TGTCGCCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCCGCCTCCTAGGTTCAAGCGA…
              |||||||||||||||||||||||||||||||||||||||||||||||||||| | |-|||||||||| |||||||||||||||||||||||| |||||||||| ||| ||||||  |||||| || |
chr2:32372685 TGTTTTGTTTGTTTGTTTGTTTGTTTGTTTGTTTGAGACGGAGTCTCACTC-TGTCGCCCAGACTGGAGTGCAGTGGTGCGATCTCAGCTCACTGCAAGCTCTGCCTCCCGGGTTCACGCCA…
                                                                                                                                *rs190625206
              | | |       |          ||  ||| ||||||||||||||||||||-|||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
chr2:32390259 TATATGTAGGACTCAAGGCAAGTTTCTTTTTTTTTGAGACGGAGTCTCGCGC-TGTCGCCCAGGCTGGAGTGTAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCTAGGTTCAAGCGA…

              |||||| | | ||| ||| ||| ||| ||| ||||||| | |||  | | | ||| ||| | |||||||||||| |||   ||||| ||||||| |||||||| ||||||  ||||||||||
chr2:32396748 TGttttctct-tttatttattt-ttt-tttttttGAGATGAAGTTACACTCTTGTTGCCTAAGCTGGAGTGCAATGGCATGATCTTGGCTCACCGCAACCTCTGCCTCCCGGGTTCAAGCGA…
                                                                                                  *rs189944254

…TTCTCCTGCCTCAGCCTCCAGAGTAGCTGGGATTACAGGCGCGTGCCACTACGCCCGGCTAA-TTTTT-GCATTTTTAGTAGAGATGGGGGTTTCACCATGTTGGTCAGGCTGAT 455
 ||||||||||||||||||  |||||||||| |||||||||   |||| ||||| ||||||-||||| | | |||||||||||| |||||||||| |||| | |||| || |
…TTCTCCTGCCTCAGCCTCCCTAGTAGCTGGGACTACAGGCGCCCACCACCACGCCTGGCTAA-TTTTTTGTACTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGT 32372918

 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||-|||||-|||||||||||||  |||||||||||||||||| ||||||| |
…TTCTCCTGCCTCAGCCTCCAGAGTAGCTGGGATTACAGGCGCGTGCCACTACGCCCGGCTAA-TTTTT-GCATTTTTAGTAGAGCCGGGGGTTTCACCATGTTGGCCAGGCTGGT 32390491

 ||||||||||||||||||  ||||||||||||||||||  | ||||| |||||||||||| |||||-| |||||||||||||||||||||||||||||||||||||||
…TTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATGGGCCACCACGCCCGGCTAATTTTTT-GTATTTTTAGTAGAGATGGGGGTTTCACCATGTTGGTCAGGCTGAT 32396979
```

This potential FoSTeS x 2 mechanism was rendered by searching for the 146bp insertion +/- 1 bp on either side using the UCSC "Short Match" within 200 kb of the CNV, and the middle genomic sequence was the only result. Red bases do not match reference sequence, despite being confirmed by both forward and reverse sequencing (of the same PCR product).

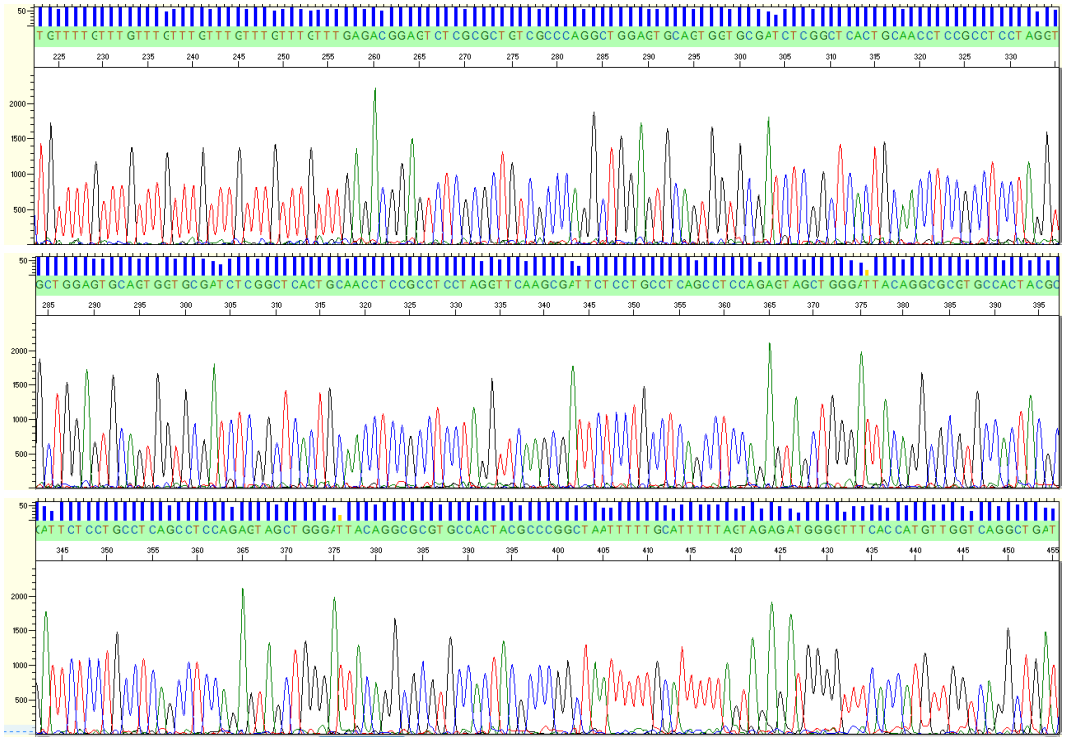


Figure S3, cont'd.

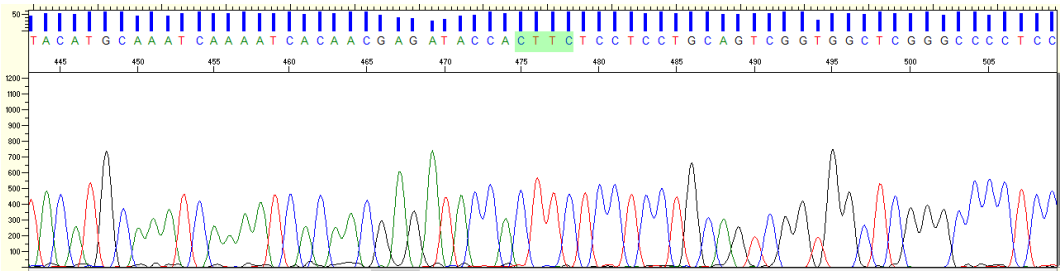
Ruhr 28

Reverse complement sequence

```
chr2:32288676 GAGGGGCCCGAGCCACCGACTGCAGGAGGAGAAGGGGTTGTGCTCCTGGCCGAGGAAGGAGAAA 32288739
              ||||||||||||||||||||||||||||||||||| |||   |   |   ||   |
Ruhr 28   508 GAGGGGCCCGAGCCACCGACTGCAGGAGGAGAAGTGGTATCTCGTTGTGATTTTGATTTGCATG 445
                        |   |  ||  |||    ||||||||||||||||||||||||||||||
chr2:32305164 TTTTAAAATAATAGTCATCCTAAAGGGTATGAAGTGGTATCTCGTTGTGATTTTGATTTGCATG 32305227
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

SPAST transcription start site underlined.



Ruhr 29

Forward sequence

```
Chr2:32348551 AGTCCCAGTCAAACCCCAGCCAGTTATTTTGTGGATACTGGCAAACTAAAGTTTATATGAAAAGGCAAAAGACCTAGAACAGCCA 32348635
              ||||||||||||||||||||||||||||||||||||||||||| | ||   |  | |        |   |  |||
Ruhr29    296 AGTCCCAGTCAAACCCCAGCCAGTTATTTTGTGGATACTGGCACAATAGTCATCCTTTCCCTTTCTGATCTGCAGAGAGAGTTGT 380
                |  ||       |     | |                   ||||||||||||||||||||||||||||||||||||||||
Chr2:33632481 TATTTTAGGGTGTATACCTATGGGTGGAACACTTCATTTTATTCAATAGTCATCCTTTCCCTTTCTGATCTGCAGAGAGAGTTGT 33632565
```

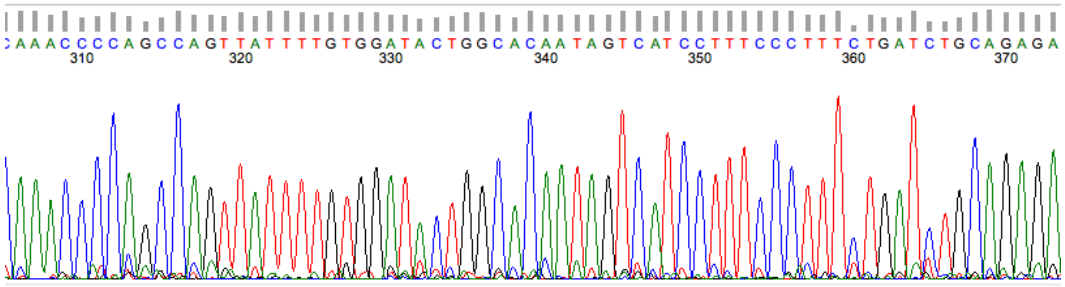


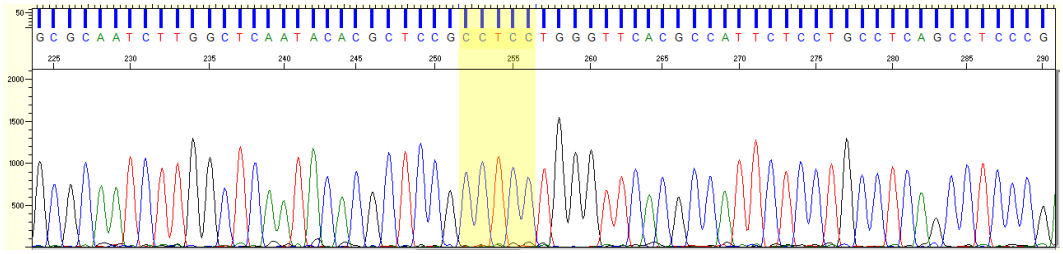
Figure S3, cont'd.

Ruhr 31

Forward sequence

```
chr2:32350903 GCCCGGGTTGGAGTGCAGTGGCGCAATCTTGGCTCAATACACGCTCCGCCTCCCGGGTTCACGCCATTTTCCTGCCTCAGCCTCCTGAGTAGCTGGGACGA 32351003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| |||||||||||||||| ||||||||||||
Ruhr 31   204 GCCCGGGTTGGAGTGCAGTGGCGCAATCTTGGCTCAATACACGCTCCGCCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTA 304
              ||||||| |||||||||||||   | |||| |||||| | || ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Chr2:32360238 GCCCGGGCTGGAGTGCAGTGGTACCATCTCGGCTCACTGCAAGCTCCACCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTA 32360338
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).



Ruhr 32

Breakpoint PCR failed to yield a product.

Park 1

Forward sequence

```
Chr2:32338700 TCTCGGCTTACTGCAACCTCCGCCTCCTGAGTTCAAGTGATTCTCCTGCCTCAGCCTCCGAAGTAGC 32338766
              ||||||||||||||||||||||||||||||||||| ||||||| ||||||| |||||  ||||||
Park_1    741 TCTCGGCTTACTGCAACCTCCGCCTCCTGAGTTCAAGCGATTCTCATGCCTCAACCTCCCGAGTAGC 807
              |||||||| |||| |||||||||||||| | ||||||||||||||||||||||||||||||||||
Chr2:32350358 TCTCGGCTCACTGAAACCTCCGCCTCCAGGGTTCAAGCGATTCTCATGCCTCAACCTCCCGAGTAGC 32350424
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

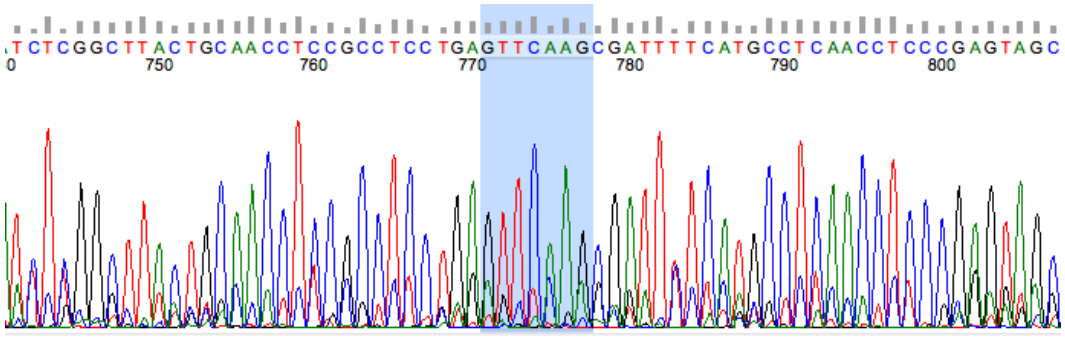


Figure S3, cont'd.

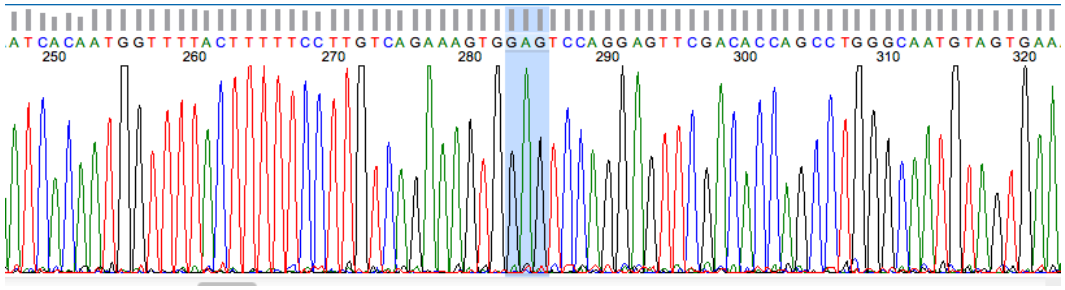
Park_2

Forward sequence

```
Chr2:32339676 AATCACAATGGTTTTACTTTTTCCTTGTCAGAAAGTGGAGCTGTTCCAAAAAGAAAAGACCCCTTAACACACACTA 32339751
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Park_2    257 AATCACAATGGTTTTACTTTTTCCTTGTCAGAAAGTGGAGTCCAGGAGTTCGACACCAGCCTGGGCAATGTAGTGA 312
              || |               |  |        | |||||||||||||||||||||||||||||||||||||||
Chr2:32347756 CCCAGCACTTTCAGAGGCCAAGGCAGGAAGATTGCTTGAGTCCAGGAGTTCGACACCAGCCTGGGCAATGTAGTGA 32347831
                                        *rs12611678
```

At underlined bases marked with "*", a known SNP exists that matches the patient sequence, potentially explaining the difference between the patient and reference genomes at that position. No other known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

The first nucleotide of *SPAST* exon 5 is underlined in red.



Park_3

Forward sequence

```
Chr2:32351019 ACGCCCGGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCATCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGAT 32351104
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| | |||| |||||||||||||| || |||| ||||||  |
Park_3    315 ACGCCCGGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTTGAACTCCCGACCTCAGGTG 400
              ||||||||||||--|||||||| |||||||||| || |||||||||||||||||||| |||||||||||||||||||||||||
Chr2:32359920 ACGCCCGGCTAA--TTTTTGTAGTTTTAGTAGACACAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCCGACCTCAGGTG 32360003
```

It is unclear whether the underlined base represents complexity or a novel SNP.

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

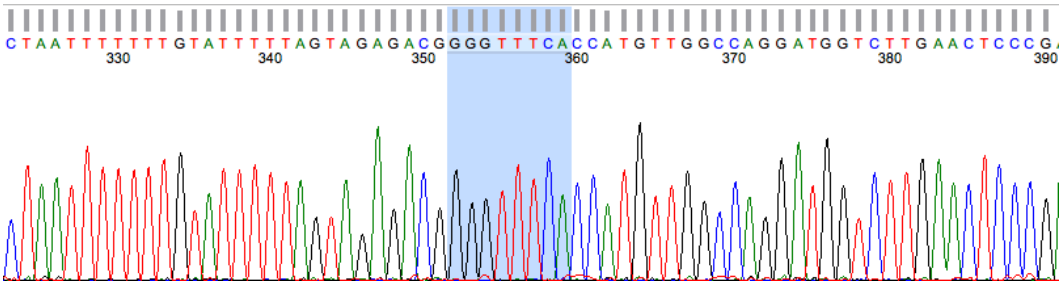


Figure S3, cont'd.

Park 4

Forward sequence

```
Chr2:32321521 ATTGGAGCAGTTTCACAGATGCTGTTTACTGTTATACTGTATGTGTCTATGACTCCTCCTCCAAGAAAAAAAAAAT 32321596
                                                  |          ||              | ||     |
Park_4     520 ATTGGAGCAGTTTCACAGATGCTGTTTACTGTTATACTGCACTCCAGCCTGGGCAGAAAGAGCAAAACTCCATCTC 595
                   ||       |  |||          |   |     |||||||||||||||||||||||||||||||||||||||||||
Chr2:32422054 GGGAGACAGAGGTTGCAGTGAGCCGAGATGGCACCACTGCACTCCAGCCTGGGCAGAAAGAGCAAAACTCCATCTC 32422129
```

No known SNPs underlie any differences (dbSNP137 via UCSC Genome Browser).

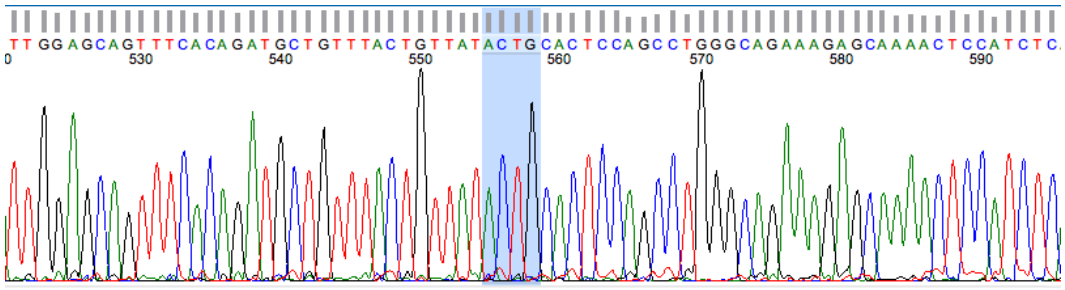


Figure S3, cont'd.

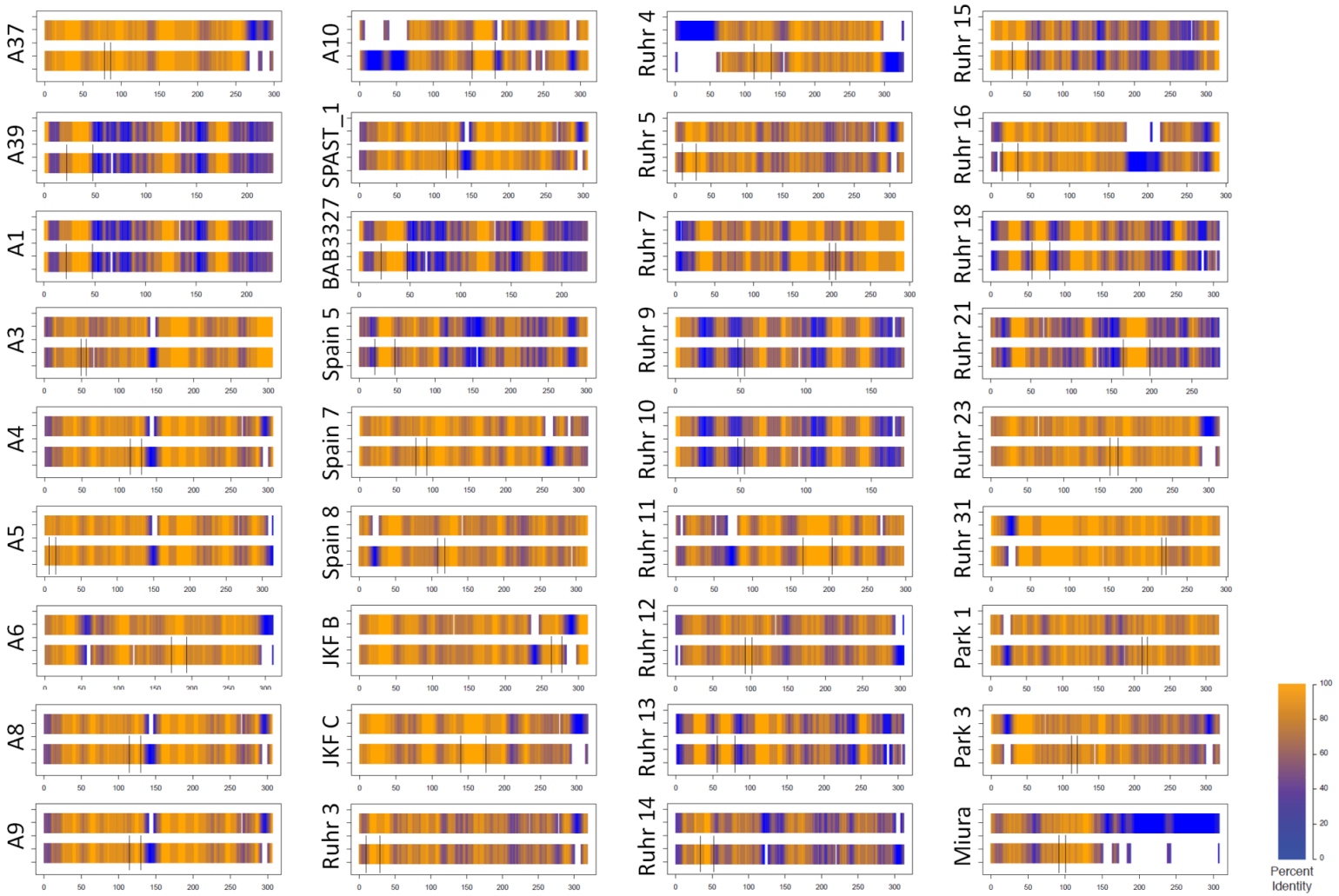


Figure S4. Breakpoint *Alu* pairs aligned for each non-complex, *Alu*-mediated CNV. The telomeric *Alu* is on top in each pairing. Heat map shading depicts a 10 bp sliding window of identity. Microhomology/crossover spots are flanked by black lines; these occupy various positions within the *Alu* pairs.

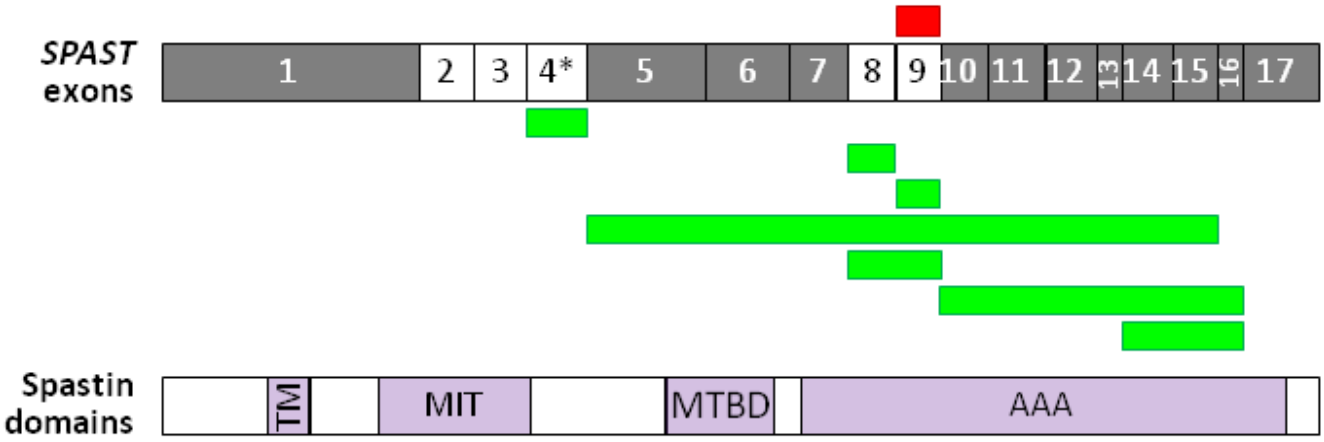


Figure S5. Pathogenic in-frame deletions and duplication of *SPAST* exons. Deletions (green) of exons 4, 8, and 9 of *SPAST*, predicted to be in-frame, in SPG4 patients suggests that these exons are indispensable. A pathogenic, predicted in-frame duplication (red) of exon 9 (24 amino acids) reinforces the sensitivity to mutation of spastin's AAA (ATPase associated with various cellular activities) domain.[1] Reported, predicted in-frame multi-exon deletions (green) are also shown. Shaded *SPAST* exons are predicted to alter reading frame when deleted or duplicated whereas unshaded exons are multiples of 3 nucleotides and thus are predicted to preserve reading frame. Data are drawn from all published *SPAST* CNVs whether breakpoint sequences were assessed or not.[2-6] Domain map from [1]. TM, transmembrane; MIT (microtubule interacting and trafficking); MTBD (microtubule binding domain); *, alternatively spliced exon.

Ruhr 9, Ruhr 10 - Haplotype in 1000 Genomes: 852/1984 (42.9%)

SNP	rs13017189	rs660393	rs572946	rs13000753	rs34460762	rs13024004	rs652978	rs201350097	rs62142107	rs6758024	rs10084410	rs78871913	rs10084280	rs17011801	rs113080319
Position (chr2)	32281280	32281509	32281510	32282297	32282340	32282362	32286411	32286411	32294435	32294486	32296099	32296134	32297984	32298072	32298273
Ref/minor allele; MAF	A/G; 47.523%	C/A; 42.811%	T/G; 2.657%	T/C; 45.263%	G/C; 10.073%	A/G; 45.359%	A/T; 47.965%	-/T; 42.883%	T/A; 5.637%	G/C; 44.538%	A/G; 3.972%	G/C; 1.218%	G/A; 5.680%	A/T; 32.088%	T/A; 1.330%
Ruhr 9	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt
Ruhr 10	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt

Ruhr 3, Ruhr 5 - Haplotype in 1000 Genomes: 1049/1984 (52.9%)

SNP	rs115440025	rs112067196	rs3769603	rs6725068	rs3830711
Position (chr2)	32360726	32360852	32360887	32360930	32362084
Ref/minor allele; MAF	T/C; 1.131%	A/G; 1.330%	C/A; 43.230%	C/T; 3.715%	-/ATAT; 36.239%
Ruhr 3	Wt	Wt	Wt	Wt	Wt
Ruhr 5	Wt	Wt	Wt	Wt	Wt

Ruhr 13, Ruhr 18 - Haplotype in 1000 Genomes: 102/1984 (5.1%)

dbSNP 138	rs12617289	rs12617290	rs7574000	rs116766118	rs13022197	rs60386495	rs150295812	rs11681731	rs6716179	rs72796871	rs72863940	rs75497994	rs139405008	rs12465351	rs7581432	rs7588883
Position (chr2)	32341399	32341408	32342888	32342989	32343167	32345123	32345168	32345325	32393076	32393157	32394674	32394704	32394747	32394905	32395315	32397267
Ref/minor allele; MAF	A/G; 46.481%	A/G; 45.480%	A/C; 30.769%	T/G; 1.436%	T/A; 44.401%	C/T; 10.046%	C/G; 1.377%	A/T; 46.460%	C/A; 50.425%	G/A; 4.003%	C/T; 45.163%	C/T; 5.440%	C/T; 1.148%	G/C; 50.601%	A/G; 50.397%	A/G; 52.168%
Ruhr 13	G	G	Wt	Wt	A	T	Wt	T	A	Wt	T	T	Wt	C	G	G
Ruhr 18	G	G	Wt	Wt	A	T	Wt	T	A	Wt	T	T	Wt	C	G	G

A1, A39, BAB 3327 - Haplotype in 1000 Genomes: 709/1984 (35.7%)

dbSNP 138	rs7572964	rs11694630	rs148193979	rs11678119	rs72796867	rs2280973	rs13425665	rs72097810	rs4952209	rs10208246	rs13429391
Position (chr2)	32372096	32376201	32376274	32376303	32376367	32378604	32385711	32385896	32386033	32386659	32386803
Ref/minor allele; MAF	A/T; 41.900%	G/A; 42.524%	A/C; 4.499%	T/C; 44.248%	T/C; 1.131%	C/A; 45.317%	C/A; 1.690%	AAG/-; 49.725%	G/C; 51.645%	G/A; 41.323%	C/T; 8.265%
A1	T	A	Wt	C	Wt	Wt	Wt	-	C	Wt	Wt
A39	T	A	Wt	C	Wt	Wt	Wt	-	C	Wt	Wt
BAB 3327	T	A	Wt	C	Wt	Wt	Wt	-	C	Wt	Wt

A4, A8, A9, SPAST 1 - Haplotype in 1000 Genomes: 895/1984 (45.1%)

dbSNP 138	rs2886393	rs7572964	rs11694630	rs148193979	rs11678119	rs72796867	rs12712326	rs111261438	rs13425665	rs72097810
Position (chr2)	32371562	32372096	32376201	32376274	32376303	32376367	32384194	32384269	32385711	32385898
Ref/minor allele; MAF	G/C; 43.046%	A/T; 41.900%	G/A; 42.524%	A/C; 4.499%	T/C; 44.248%	T/C; 1.131%	C/G; 9.817%	G/A; 2.570%	C/A; 1.690%	AAG/-; 49.725%
A4	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt
A8	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt
A9	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt	Wt
SPAST_1	*	*	*	*	*	*	*	*	*	*

Figure S6. SNPs flanking identical CNVs found in multiple individuals. To identify whether identical CNVs occurred on the same haplotype (i.e. no specific evidence for recurrent CNV formation) or different haplotype (i.e. suggests recurrence), common SNPs

(dbSNP 138 allele frequency ≥1%) flanking the CNVs were genotyped. No flanking SNPs differ among individuals sharing the same CNV, thus failing to provide evidence for recurrence. Note that some SNPs differ between individuals A1/A39/BAB3327 and individuals A4/A8/A9, indicating that *SPAST* CNVs have occurred on different haplotypes, a positive control of our method. SNPs are listed in genomic order. The empty space in each box indicates the location of the CNV. Reference alleles are listed as "Wt" and shaded in lavender. Non-reference alleles are shaded in blue. MAF, minor allele frequency (dbSNP 138). *SPAST 1 SNPs could not be analyzed owing to failure at the PCR stage. PCR and raw sequencing data are not shown.

Supplemental Tables

Breakpoint PCR Primers[a]	
Primer name	**Sequence (5' to 3')**
A1 F1	TTTTTCTATTGCCTGGACTCTGTGAACCTATC
A1 R2	CTATCCACCATGTGAATAACTGCCTCAATCT
A2 F1	GGTGAAGAATATGTCTGCCAGTGAGGTATAGT
A2 R1	GGTCTCGTGAAAATTATATTCTGGTAGGGAAG
A3 F1	TATTCTTCTACACTTTGGCCTTTAGGAAGAGG
A3 R1	CAAATGAATGAGTGTAGTGTGTGGATGTGT
A4 F1, A8 F1, A9 F1	TATTCTTCTACACTTTGGCCTTTAGGAAGAGG
A4 R1, A8 R1, A9 R1	GTTGAGCTGGAACCACATAGTCCTTTAATATG
A5 F1	TATTCTTCTACACTTTGGCCTTTAGGAAGAGG
A5 R1	GAATAGTTCCAGAAGGAGAGTTGTTCAAGAGA
A6 F1	GTAGATCATTGTACTTGGTTTTGCCCTTCAAC
A6 R1	TTATTACACCAAAGAGCAGTATCTTCCAGGAC
A10 F1	TGATGCTCTCAGGTATGCACTCAGATTATTAG
A10 R1	GACTACTCAAGATACCTATCCCTTCCCCTACA
BAB 3200 F1	AGCAATGCCTAAGAAACTCCTATAACTCCTGT
BAB 3200 R1	AAACCAACTCCTCGGTGGTAAAAATGAAAG
BAB 3327 F1	TTTTTCTATTGCCTGGACTCTGTGAACCTATC
BAB 3327 R2	CTATCCACCATGTGAATAACTGCCTCAATCT
BAB 5112 F1	CTCTGTTGCATAGAGAAGATGCAACCAGTTTT
BAB 5112 R1	GTTTGTGTACTGCCTGGCTGCTATCTATTACT
SPAST_1 F1	ATATAGTGCTTATGGGGTTTTCTTTGGTGGGG
SPAST 1 R1	GCTGGAACCACATAGTCCTTTAATATGATTAGCC
SPAST_3 F1	TTATTGGTGAGGACTTCTGTTTAGGGGCTTCA
SPAST 3 R1	AATGGCTATCACAGGTTTTCAGTCTTCCAAGA
Spain 1 F1	GACTCTCCATTTCTTAGGTTCTCTGACATCTG
Spain 1 R1, Spain 7 R1	TTTGCTCACAGAGTTATGACCATTCTGCATAC
Spain 2 F1	CAGAGTACTTGTAATGGGTGCAACTAATAGGC
Spain 2 R1	TCCTTTTCTCTATTCCAACTCTCAGCACAT
Spain 3 F1	CCACACTTGTAATCTCATACAGCAGTTCTCAT
Spain 3 R1	GTTTACCATCTAATCACAGTGAGCATCTGAAC
Spain 4 F1	CTAACTGAGGTCTTGTTTCTTAGTGGAACAGC
Spain 4 R1	CCAGAGTCTTCCTCTTCAAAGACACCTATTAT
Spain 5 R1	CAGAGCTGTACAAATTAACTGGCCTCCAAT
Spain 6 F1, Spain 8 F1	GGTGAACTATGAGAGAATTAATCAGGGGCTAT
Spain 6 R1	TATACGCAGGTTAAGCTTGTCTTCAGCAAATC
Spain 7 F1	CAGCGAGAGAAACTTGATTTTATCAGTACACC
Spain 8 R1	GCTGCCCATAGCTTTAAAGGAAACACATAATC
Spain 9 F1	TTAACACACACTAGTAATTCACTGCCTCGTTC
Spain 9 R1	TGAAGGATGAATATCTAGAAAGGTCCAAGAGG
JKF B F1	ATTACCAGTAACACCTCCCGTAGTGAAATAGG
JKF B R1	TATCTGTGTTAACAGTTCACTGTCCCCTCAAC
JKF C F1	CACACCGTCTTCAAAATGTAAAGTTTGGTC
JKF C R1	GGAAACAGAGCACTCACGTATTTTGAAGTT
Ruhr 3 F1, Ruhr 5 F1	TTCATTCATAGTGTCTGTCTCTACTGCCAACT
Ruhr 3 R1, Ruhr 5 R1	AGTTCCTCCCAAAAGCTTTCTCAGGATAGAT
Ruhr 4 F1, Spain 5 F1	GCGCTTCAGGGGCTAACAATATACTTATTTC
Ruhr 4 R1	CAGTGATATCCCCCTTTGCTTTACTATATTCC
Ruhr 6 F1	TCTTACAGTATAGTCAGTGGCAGTTTCCCTCT
Ruhr 6 R1	GCTGTTCCACTAAGAAACAAGACCTCAGTTAG
Ruhr 7 F1	CAAGGAAGAATATTACTGGTCCCTTTTGTGAG

Table S2. Primer sequences.

Ruhr 7 R1, Ruhr 21 R1	GTTGAGCTGGAACCACATAGTCCTTTAATATG
Ruhr 9 F1, Ruhr 10 F1, Ruhr 28 F1	AAGAGCACTCGTTTCCAGTTACTTTCTATTCG
Ruhr 9 R1, Ruhr 10 R1	GATACTGAAAGATTTAAGTCACCAGCCTGAGA
Ruhr 11 F1	AATAACCATCCTAATGGGCTTGATATGGTC
Ruhr 11 R1	AGTGAAATAACTGTGACCCAGAAGTGAGTCTT
Ruhr 12 F1	CAAGACTGTTGTTTAAAATGTGCTCCCCTA
Ruhr 12 R1	CCTTAATTTCTCCTTTTATCCTACCTCCCAAC
Ruhr 13 F1, Ruhr 18 F1	CCTCTTGGACCTTTCTAGATATTCATCCTTCA
Ruhr 13 R1	AGGGCTGTGAGTTAATGAAGGTGTATAATCTG
Ruhr 14 F1	GGTCCTATCCGAGGTAGGTATACAAGAGCTTA
Ruhr 14 R1	CCAGAGTCTTCCTCTTCAAAGACACCTATTAT
Ruhr 15 F1	TTCATTCATAGTGTCTGTCTCTACTGCCAACT
Ruhr 15 R1	CTATGAAACAAGAGCCCAGTGAAATGTAAGAG
Ruhr 16 F1	GGCCACTCTAATTTGCTGACTGACTTAACAT
Ruhr 16 R1	CCCTTGTTAGAATCATGACAAAACTCTAGG
Ruhr 17 R1	GTTTGGCATTCCAATCACCTGTAGTTAAGAC
Ruhr 18 R1	CCAGATCTGTCTAACCCAAGCTCTTATTCAT
Ruhr 21 F1	ATCTATCCTGAGAAAGCTTTTGGGAGGAACT
Ruhr 23 F1, Ruhr 27 F1	GTGAACTATTGTTTTTGGGCAGTATGCAAG
Ruhr 23 R1	GGACCAAATAAGATAAACCGAATCAGCTACAG
Ruhr 26 F1	GTTATGTTGCGTCCTTTTGTTTTCCCTATC
Ruhr 26 R1	TATCTGTGTTAACAGTTCACTGTCCCCTCAAC
Ruhr 27 R1	GAATAGTTCCAGAAGGAGAGTTGTTCAAGAGA
Ruhr 28 R2	CTTTAACATTAGCTCCCAACAGTCCCTAGTTT
Ruhr 29 F2	CATCATACCACACACCTTCTACTCCTGTCCTC
Ruhr 29 R2	CAAATATGATGTAAAGCCCCCAGTCAGTCTTC
Ruhr 31 F1, Ruhr 17 F1	TAAAGTGTAGGTGAGAtgttttcagtctggag
Ruhr 31 R1	CACCAAAGTAATGAGTCATGAGGTCTGTGATA
Ruhr 32 F1	AGTGCTGATGTTATAATCCAAGCCCAAATG
Ruhr 32 R1	GGGTCAATGGACTTCCTTGTTTACATAACAG
Ruhr 32 F2	AAACTAGGGACTGTTGGGAGCTAATGTTAAAG
Ruhr 32 R2	GATTAAGGCATACTATCTCCAGAGGCTTCATT
Park 1 F1, Park 2 F1	TTGGCCACTCTAATTTGCTGACTGACTTAACA
Park 1 R1, Park 2 R1	TTTTCAGTAGCATCTTCCCAAACAGAGCTGC
Park 3 F1	CCTTAGCTCCATCATACCACACACCTTCTACT
Park 3 R1	ACTTGCGCTATTAAGAACACCGTGAAAAGTCA
Park 4 F1	TATCTGGCACCACATCTCTGAAAAGCTCCAAT
Park 4 R1	CTCTTTTCTGTTTCACTCATCAGGTCAGCAGT
Long range Ctrl F	CTAGAAGGAACCTAGATAGGTCAAGCCTCTGC
Long range Ctrl R	AGAGGAGCTTCAGATGCTGGAAAGCTAGAG
cDNA PCR primers	
Primer name	**Sequence (5' to 3')**
SPAST exon 16 F1	CTAAAACCAGAACAGGTGAAGAA
SLC30A6 exon 2 R1	TCAGCTGCTACAAGTCTAAATTCC
SPAST exon 12 F1	CAGTCTGCTGGAGATGACAGA
SLC30A6 exon 7 R1	GGCAAATACAGCCAGGACTT
External pos ctrl (KDM4C exon 18) F	CTGCATCCAGTGTTCCTACG
External pos ctrl (KDM4C exon 19) R	CCTCAGGAAATGTGTCTCTGC
Haplotyping PCR primers	
Primer name	**Sequence (5' to 3')**
Ruhr 9-10 Haplo F1	TCTAGAATCTCTCGAATTGAAGGTCTGAAGAG
Ruhr 9-10 R1	GATACTGAAAGATTTAAGTCACCAGCCTGAGA
Ruhr 9-10 Haplo R1	CTGTTTAAATACCTCTCGTTCTTCCCCATATC
Ruhr 9-10 F1	AAGAGCACTCGTTTCCAGTTACTTTCTATTCG
Ruhr 3-5 Haplo F1	GCAGTACTGGCGGGGAATAATTAAGTAAAA
Ruhr 3-5 R1	AGTTCCTCCCAAAAGCTTTCTCAGGATAGAT
Ruhr 3-5 Haplo R1	GATAGGGAAAACAAAAGGACGCAACATAACAC
Ruhr 3-5 F1	TTCATTCATAGTGTCTGTCTCTACTGCCAACT

Table S2, cont'd.

Ruhr 13-18 Haplo F1	CCTACCACTCATAAGGTATTCTGGGACAGTAA
Ruhr 13 R1	AGGGCTGTGAGTTAATGAAGGTGTATAATCTG
Ruhr 13-18 Haplo R1	GAATAGTTCCAGAAGGAGAGTTGTTCAAGAGA
Ruhr 13 F1	CCTCTTGGACCTTTCTAGATATTCATCCTTCA
A1-A39-BAB3327 Haplo F1	GCTGCTGGAGCTTAAATCTTGTTCAGTTAG
A1-A39-BAB3327 R1 (A1 R2)	CTATCCACCATGTGAATAACTGCCTCAATCT
A1-A39-BAB3327 Haplo R1	GGGACCAAATAAGATAAACCGAATCAGCTAC
A1-A39-BAB3321 F1 (A1 F1)	TTTTTCTATTGCCTGGACTCTGTGAACCTATC
A4-8-9 Haplo F1	GCTGGAGCTTAAATCTTGTTCAGTTAGTGGAT
A4-8-9 R1	GTTGAGCTGGAACCACATAGTCCTTTAATATG
A4-8-9 Haplo R1	TTCAGCTCTTCCATCCCACTGTTTTATCTTAC
A4-8-9 F1	TATTCTTCTACACTTTGGCCTTTAGGAAGAGG
Haplotyping sequencing primers	
Primer name	**Sequence (5' to 3')**
Ruhr_9-10 Haplo Seq L F1	TGTCCTGAAAGTTCTTACAAGCT
Ruhr_9-10 Haplo Seq L R1	AGGAGAGAACAGCAGTGACA
Ruhr_9-10 Haplo Seq L R2	TGACCTACTCCCTGTTTCCC
Ruhr_9-10 Haplo Seq R F1	CAGAAGATGAGGATGGAGGGT
Ruhr_9-10 Haplo Seq R R1	TGACCCTTTAGCCCTTTAAAACT
Ruhr_9-10 Haplo Seq R R2	AGGGAATGGCAGTGCATGAT
Ruhr_3-5 Haplo Seq L F1	ATGTTTGATGAATGGATTGCGT
Ruhr_3-5 Haplo Seq L R1	GCAACCGTTTTAATTCTCCCC
Ruhr_3-5 Haplo Seq L R2	ACGCAATCCATTCATCAAACAT
Ruhr_3-5 Haplo Seq R F1	GAAGAGAAGGGGAGCACGAT
Ruhr_3-5 Haplo Seq R R1	TGTTGAAAATCTAGCACACAGGA
Ruhr_3-5 Haplo Seq R F2	CTGGGAATGTCTCAAGGGGT
Ruhr_13-18 Haplo L F1	CTTCCTTCTCTGAGGCCTGA
Ruhr_13-18 Haplo L R1	ATAATGCACCCACCACAACG
Ruhr_13-18 Haplo L R2	AGTGGGAGGACAGGAGTAGA
Ruhr_13-18 Haplo R R1	TCACTGGGAGAAAATATAAGCCA
Ruhr_13-18 Haplo R F1	GGCCACCAATGTAAACCTTCT
Ruhr_13-18 Haplo R F2	TGAAGGGAGTTGAAGATAAAGGG
A1-39_BAB3327 Haplo L R1	GCAGTTGAAATTGTTGAAGGGC
A1-39_BAB3327 Haplo L R2	CTTGTGTTTACGTCTCATCCCA
A1-39_BAB3327 Haplo L R3	TAGGTTCACAGAGTCCAGGC
A1-39_BAB3327 Haplo R R1	ACAGTTCAGGAAGGAGGAAGG
A1-39_BAB3327 Haplo R R2	TACCATCTCTGCTGACACCC
A_4-8-9 Haplo L R1	GCAGTTGAAATTGTTGAAGGGC
A_4-8-9 Haplo L R2	CTTGTGTTTACGTCTCATCCCA
A_4-8-9 Haplo L F1	GCTGGAGCTTAAATCTTGTTCAGTTAGTGGAT
A_4-8-9 Haplo R R1	TTCAGCTCTTCCATCCCACTGTTTTATCTTAC
A_4-8-9 Haplo R F1	CAGATGCACAAGTGACCTCT
A_4-8-9 Haplo R R2	AGAGGTCACTTGTGCATCTG

[a]Primers used exclusively for breakpoint sequencing are available upon request.

Table S2, cont'd.

STR Marker	AMEL	CSF1PO	D13S317	D16S539	D18S51	D19S433	D21S11	D2S1338	D3S1358	D5S818	D7S820	D8S1179	FGA	TH01	TPOX	vWA
Dye	R	B	G	G	Y	Y	B	G	G	R	B	B	R	G	Y	Y
Ruhr 9	X / Y	11 / 12	11 / 11	9 / 14	14 / 18	13 / 14.2	28 / 30	20 / 22	18 / 18	11 / 12	8 / 10	12 / 13	18 / 24	9.3 / 9.3	8 / 8	15 / 16
Ruhr 10	X / Y	10 / 10	11 / 12	10 / 13	13 / 18	10 / 14	30 / 30.2	17 / 19	15 / 16	11 / 13	8 / 11	12 / 17	19 / 22	7 / 8	8 / 8	16 / 16
Ruhr 3	X / Y	12 / 12	8 / 8	12 / 13	11 / 13	13 / 14	29 / 32	23 / 24	16 / 19	11 / 13	10 / 12	10 / 14	21 / 21	6 / 9.3	8 / 8	17 / 19
Ruhr 5	X / Y	10 / 11	12 / 12	12 / 13	14 / 19	14 / 15	29 / 31	23 / 24	15 / 16	12 / 13	11 / 12	11 / 13	21 / 23	7 / 7	8 / 8	16 / 18
Ruhr 13	X / Y	10 / 10	10 / 13	11 / 13	11 / 20	12 / 14	28 / 30	17 / 19	15 / 16	13 / 13	10 / 11	10 / 11	24 / 25	6 / 9	8 / 12	17 / 17
Ruhr 18	X / Y	9 / 10	11 / 13	9 / 9	13 / 17	14 / 14.2	28 / 29	17 / 23	14 / 17	11 / 14	8 / 11	9 / 14	22 / 22	9 / 9.3	8 / 11	15 / 15
A1[a]	X / Y	11 / 11	9 / 12	11 / 14	13 / 15	14 / 16	29 / 32	19 / 20	15 / 16	11 / 12	10 / 12	10 / 10	21 / 21	9.3 / 9.3	8 / 8	14 / 17
A39	X / X	10 / 11	11 / 12	11 / 12	13 / 16	15 / 16	28 / 31.2	20 / 25	15 / 16	12 / 12	8 / 12	10 / 13	22 / 24	9 / 9.3	9 / 11	16 / 17
BAB 3327[a]	X / X	10 / 11	11 / 12	9 / 11	13 / 18	14 / 16	30 / 32	19 / 19	15 / 16	11 / 12	10 / 12	10 / 13	20 / 21	7 / 9.3	8 / 9	14 / 17
A4	X / X	12 / 12	8 / 12	11 / 12	13 / 16	14 / 14	30 / 30.2	20 / 20	15 / 17	12 / 12	10 / 11	11 / 13	20 / 24	6 / 9	8 / 9	17 / 17
A8	X / X	9 / 11	10 / 12	10 / 13	12 / 12	14 / 15	27 / 30	19 / 25	14 / 18	12 / 12	10 / 13	12 / 16	21 / 21	7 / 9	8 / 8	17 / 17
A9	X / X	9 / 10	9 / 11	9 / 11	16 / 17	14 / 15.2	28 / 30	20 / 26	16 / 17	11 / 12	7 / 9	12 / 13	22.2 / 25	6 / 6	8 / 11	15 / 18
SPAST_1	X / X	11 / 12	8 / 12	11 / 13	13 / 17	13 / 14	29 / 31.2	17 / 23	16 / 19	10 / 12	10 / 11	10 / 13	19 / 21	6 / 8	8 / 8	14 / 18

[a] STR alleles consistent with a parent-child pair.

STR allele sizes are separated by a "/". If only a single STR allele was present, it is listed once on either side of the "/". Raw electrophoretic data available upon request.

Table S3. STR alleles indicate that subjects with identical CNVs are unique individuals.

References

1. Shoukier, M., Neesen, J., Sauter, S.M., Argyriou, L., Doerwald, N., Pantakani, D.V.K., and Mannan, A.U. (2009). Expansion of mutation spectrum, determination of mutation cluster regions and predictive structural classification of *SPAST* mutations in hereditary spastic paraplegia. Eur J Hum Genet 17, 187-194.

2. Boone, P.M., Liu, P., Zhang, F., Carvalho, C.M.B., Towne, C.F., Batish, S.D., and Lupski, J.R. (2011). *Alu*-specific microhomology-mediated deletion of the final exon of *SPAST* in three unrelated subjects with hereditary spastic paraplegia. Genet Med 13, 582-592.

3. Álvarez, V., Sánchez-Ferrero, E., Beetz, C., Díaz, M., Alonso, B., Corao, A.I., Gámez, J., Esteban, J., Gonzalo, J.F., Pascual-Pascual, S.I., et al. (2010). Mutational spectrum of the *SPG4* (*SPAST*) and *SPG3A* (*ATL1*) genes in Spanish patients with hereditary spastic paraplegia. BMC Neurol 10, 89.

4. Miura, S., Shibata, H., Kida, H., Noda, K., Toyama, T., Iwasaki, N., Iwaki, A., Ayabe, M., Aizawa, H., Taniwaki, T., et al. (2011). Partial *SPAST* and *DPY30* deletions in a Japanese spastic paraplegia type 4 family. Neurogenetics 12, 25-31.

5. Sulek, A., Elert, E., Rajkiewicz, M., Zdzienicka, E., Stepniak, I., Krysa, W., and Zaremba, J. (2013). Screening for the hereditary spastic paraplaegias SPG4 and SPG3A with the multiplex ligation-dependent probe amplification technique in a large population of affected individuals. Neurol Sci 34, 239-242.

6. Loureiro, J.L., Brandão, E., Ruano, L., Brandão, A.F., Lopes, A.M., Thieleke-Matos, C., Miller-Fleming, L., Cruz, V.T., Barbosa, M., Silveira, I., et al. (2013). Autosomal dominant spastic paraplegias: a review of 89 families resulting from a Portuguese survey. JAMA Neurol 70, 481-487.