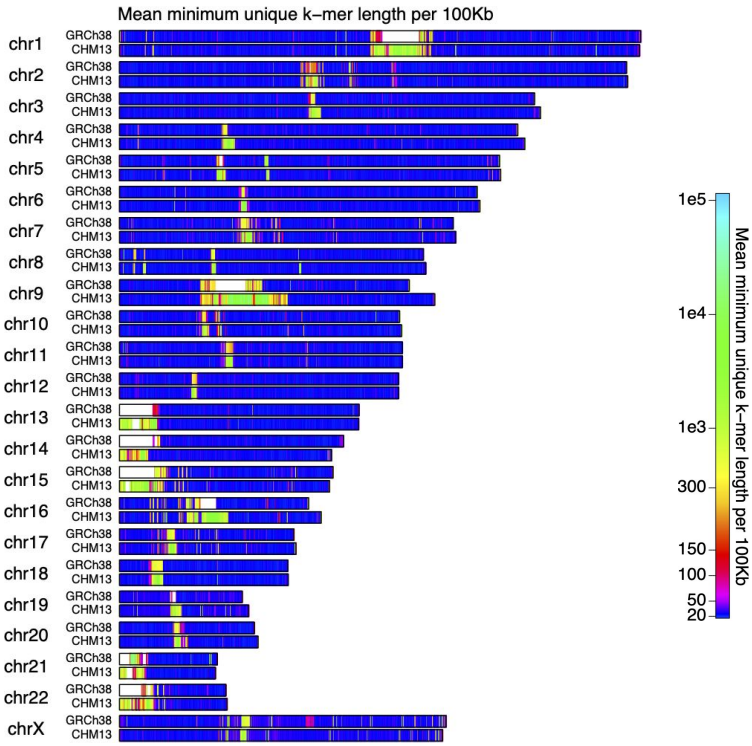
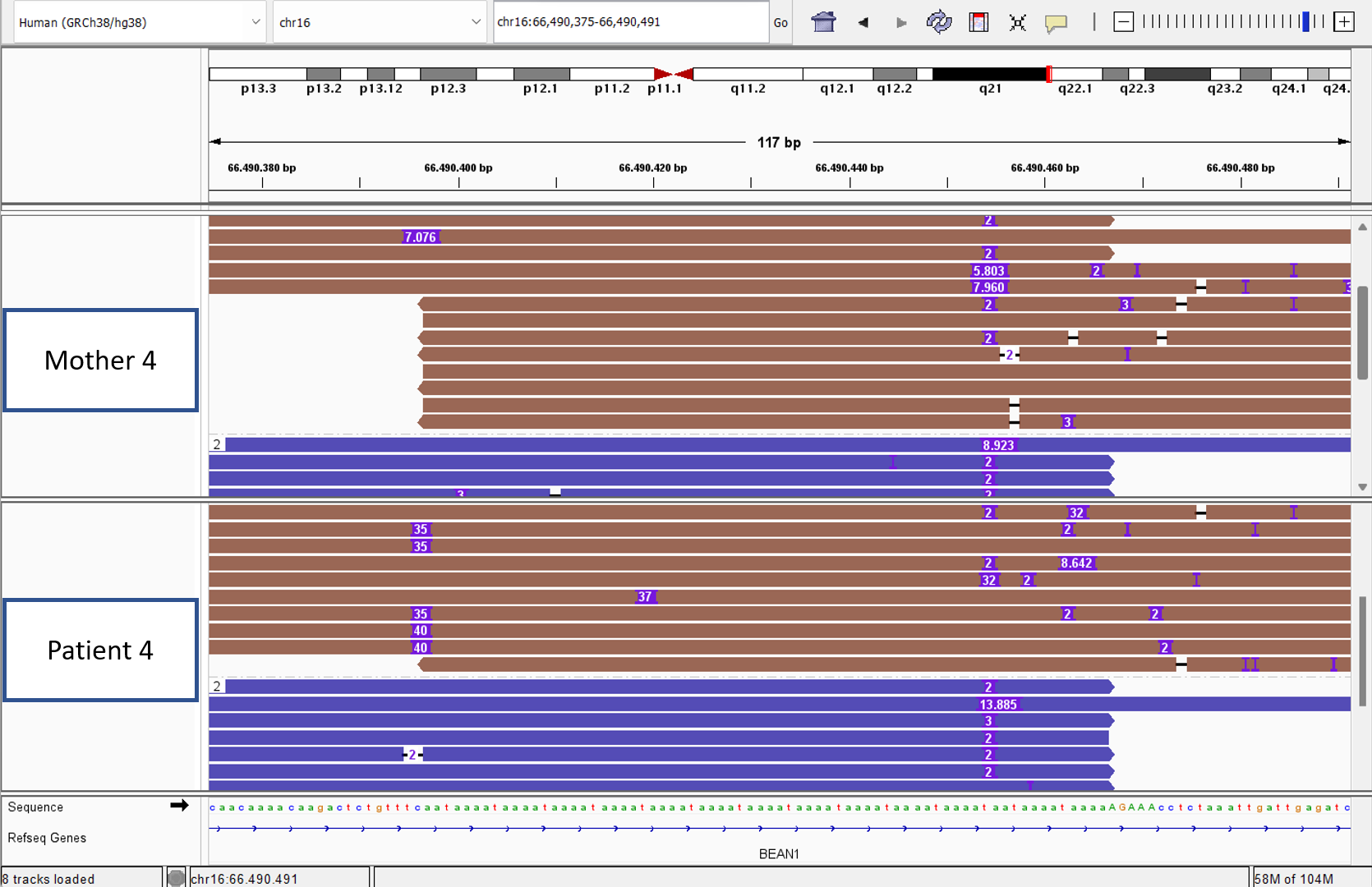
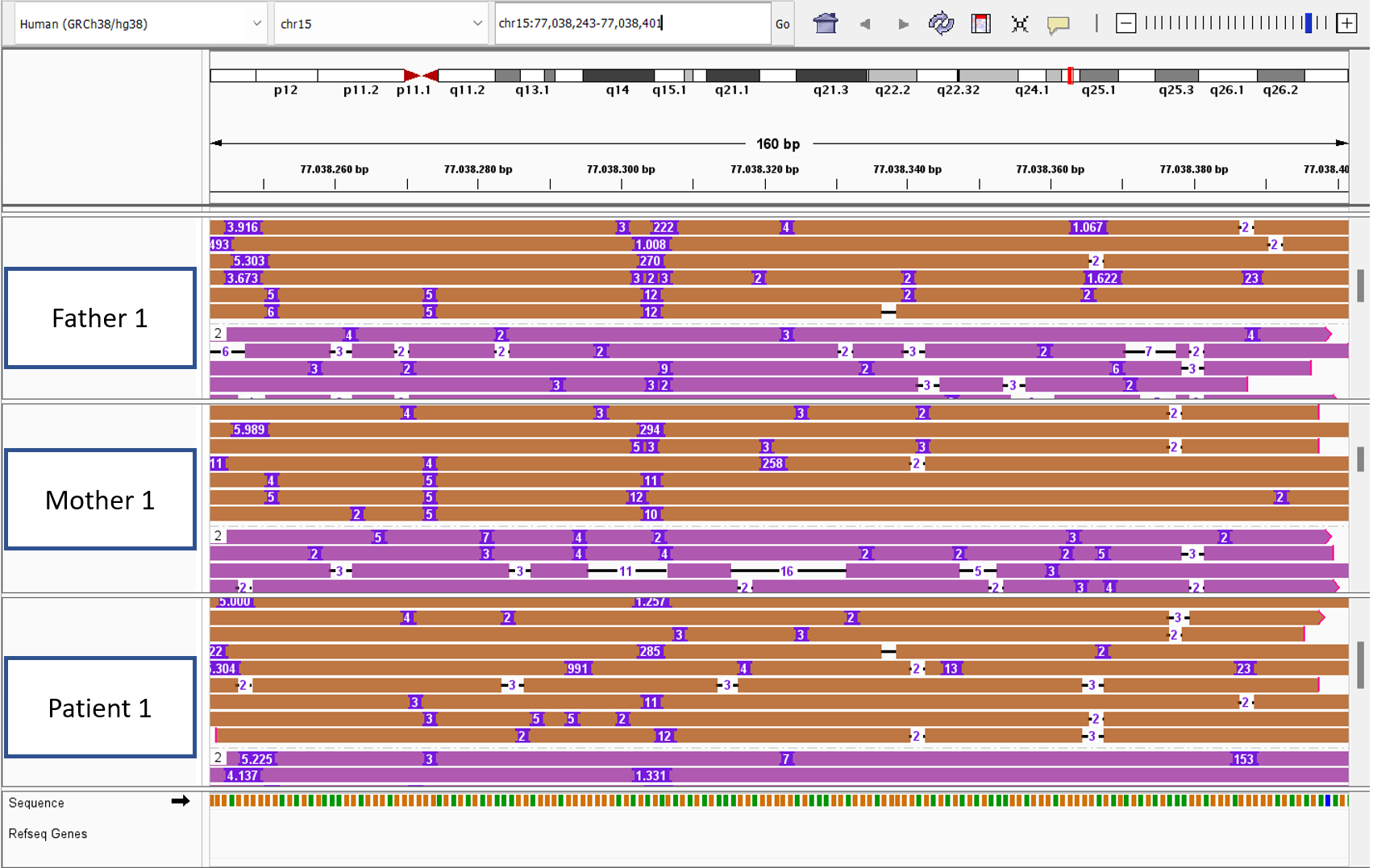
# Supplementary figures



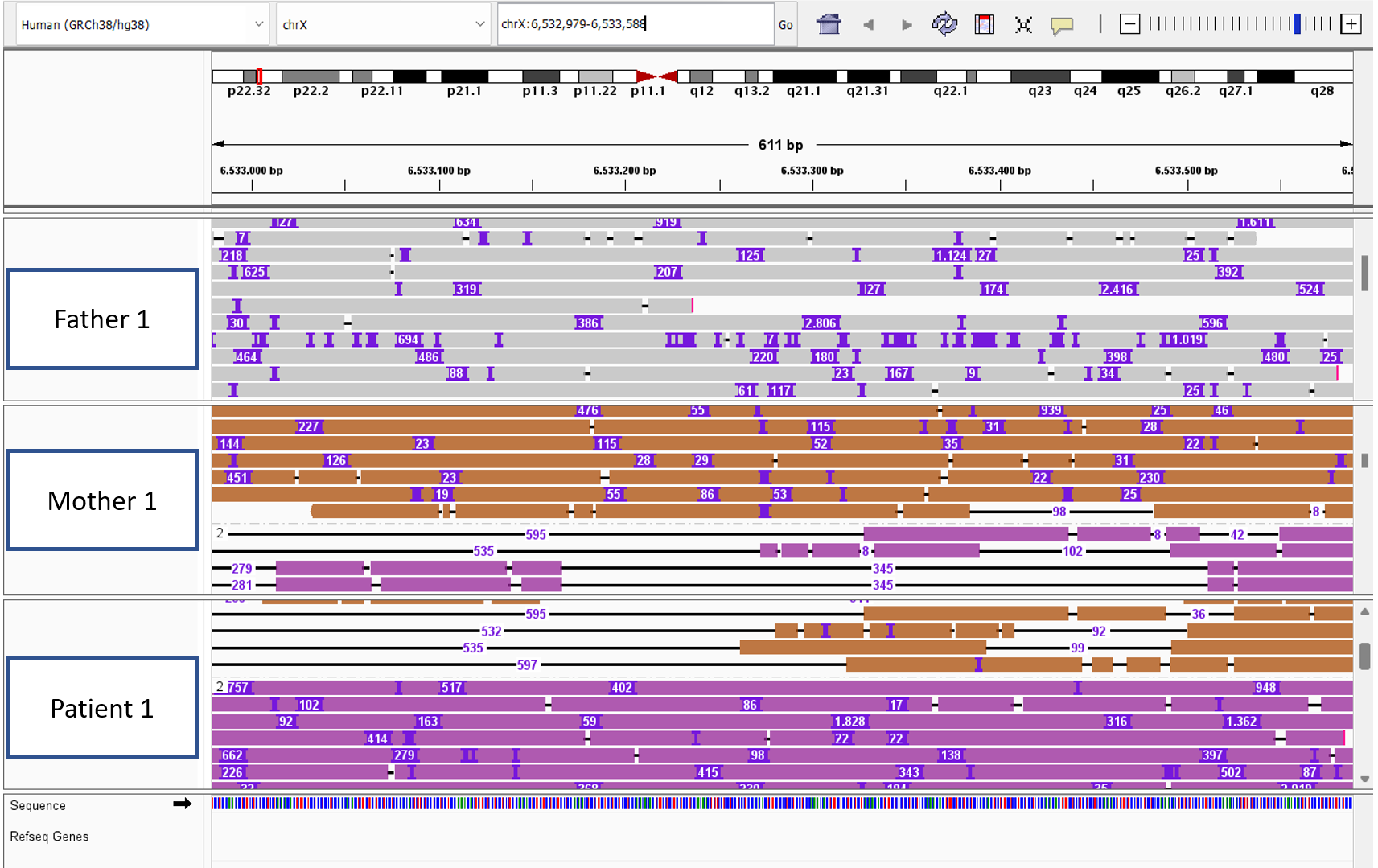
**Supplementary figure 1: Minimum unique k-mer chromosome score maps (Agenezov et al.).** This figure shows the average number of unique k-mers per 100 kbp in human reference genome 38 (GRCh38) and the telomere-to-telomere reference genome (CHM13). Regions shown in white indicate that no valid score exists, because the k-mer consists of at least one N. This figure was taken from Agenezov et al. (Fig. S.5)



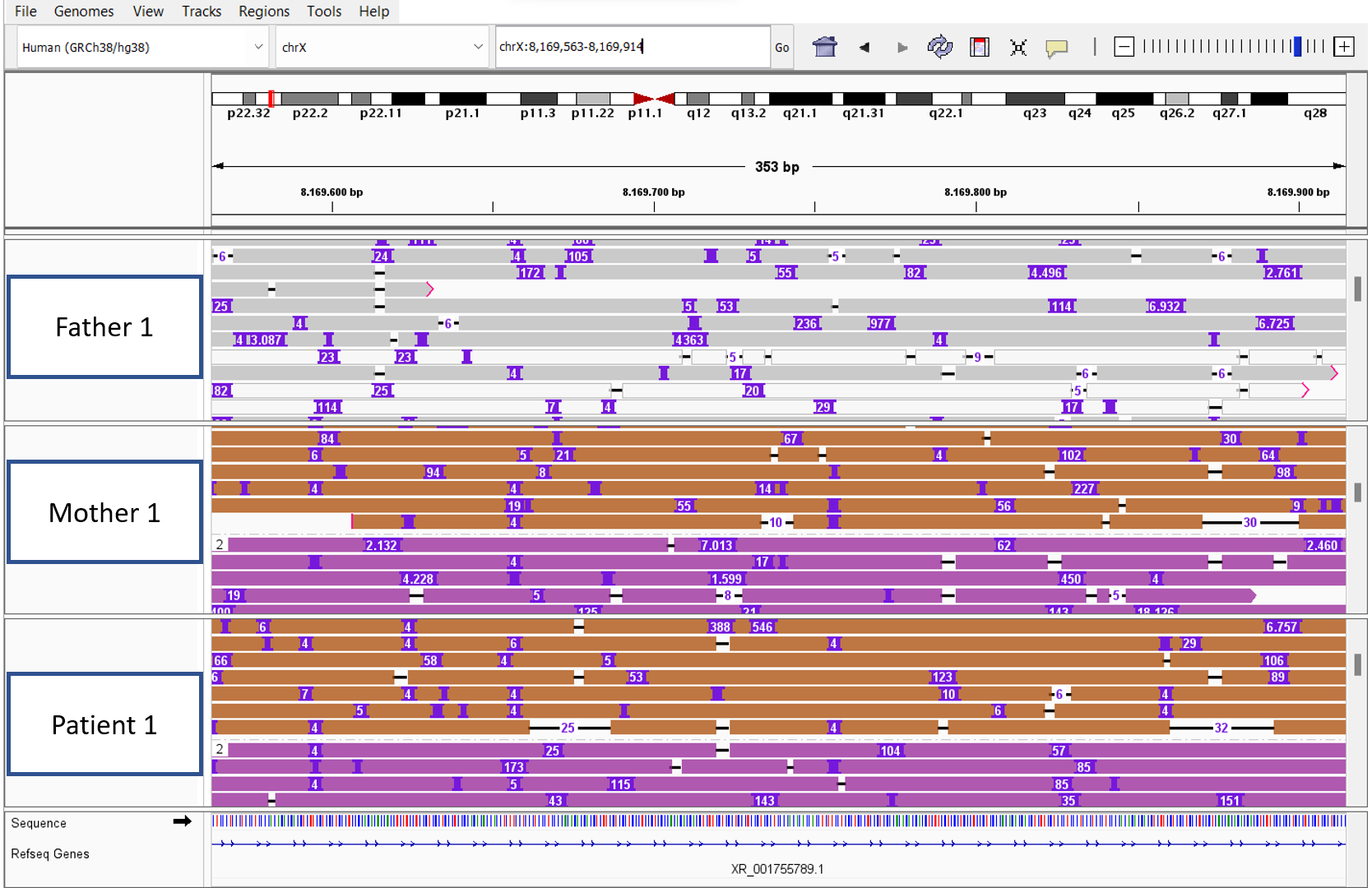
**Supplementary figure 2: Sequencing data Patient 4 and her mother in the region of the tandem-repeat, associated with Spinocerebellar ataxia 31 (SCA31).** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated with SCA31. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first track show reads from the mother of trio 17.10085, and the second track shows reads from the patient of trio 17.10085 (patient 4). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



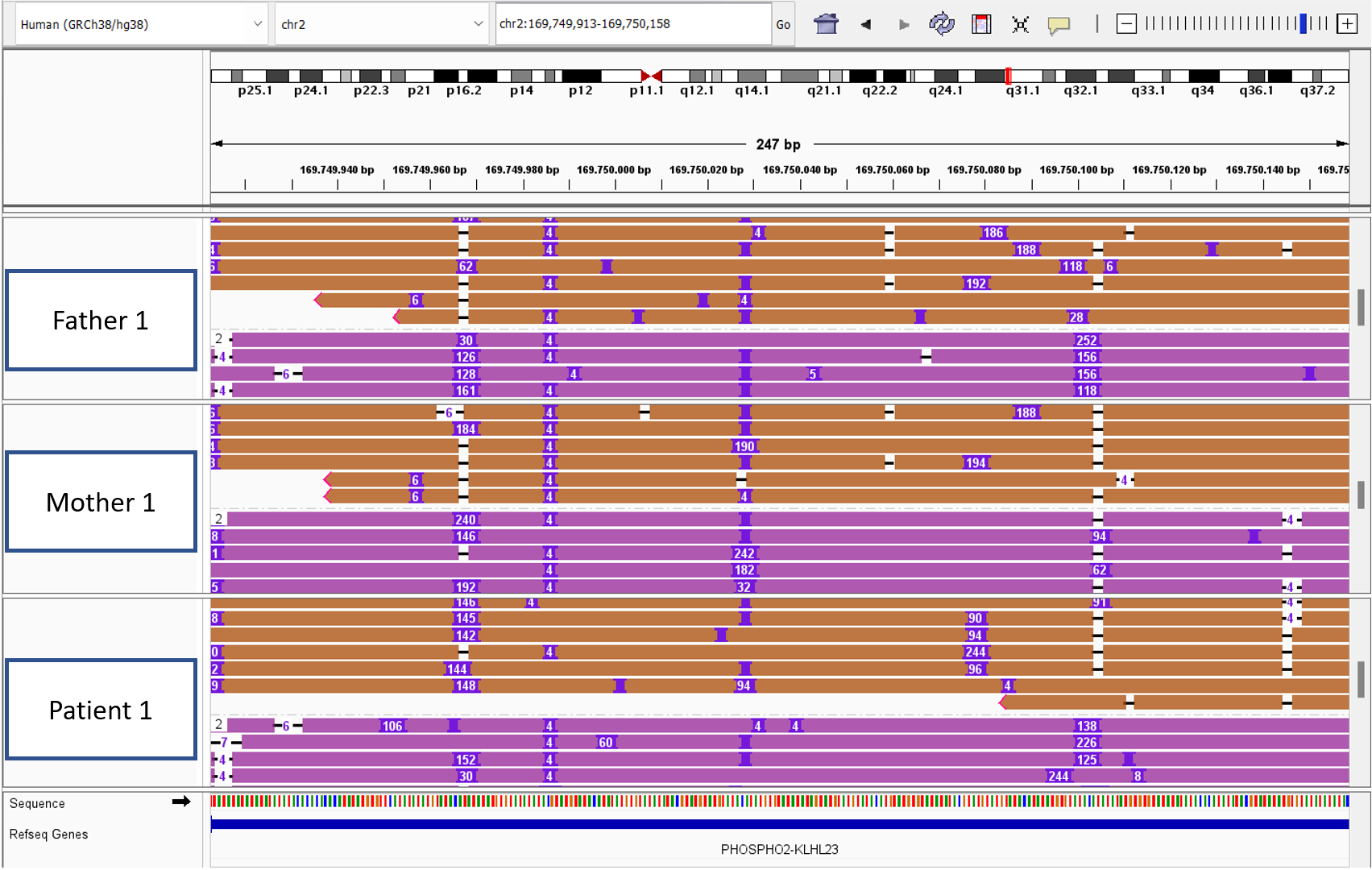
**Supplementary figure 3: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by Straglr on chromosome 15.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by Straglr. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



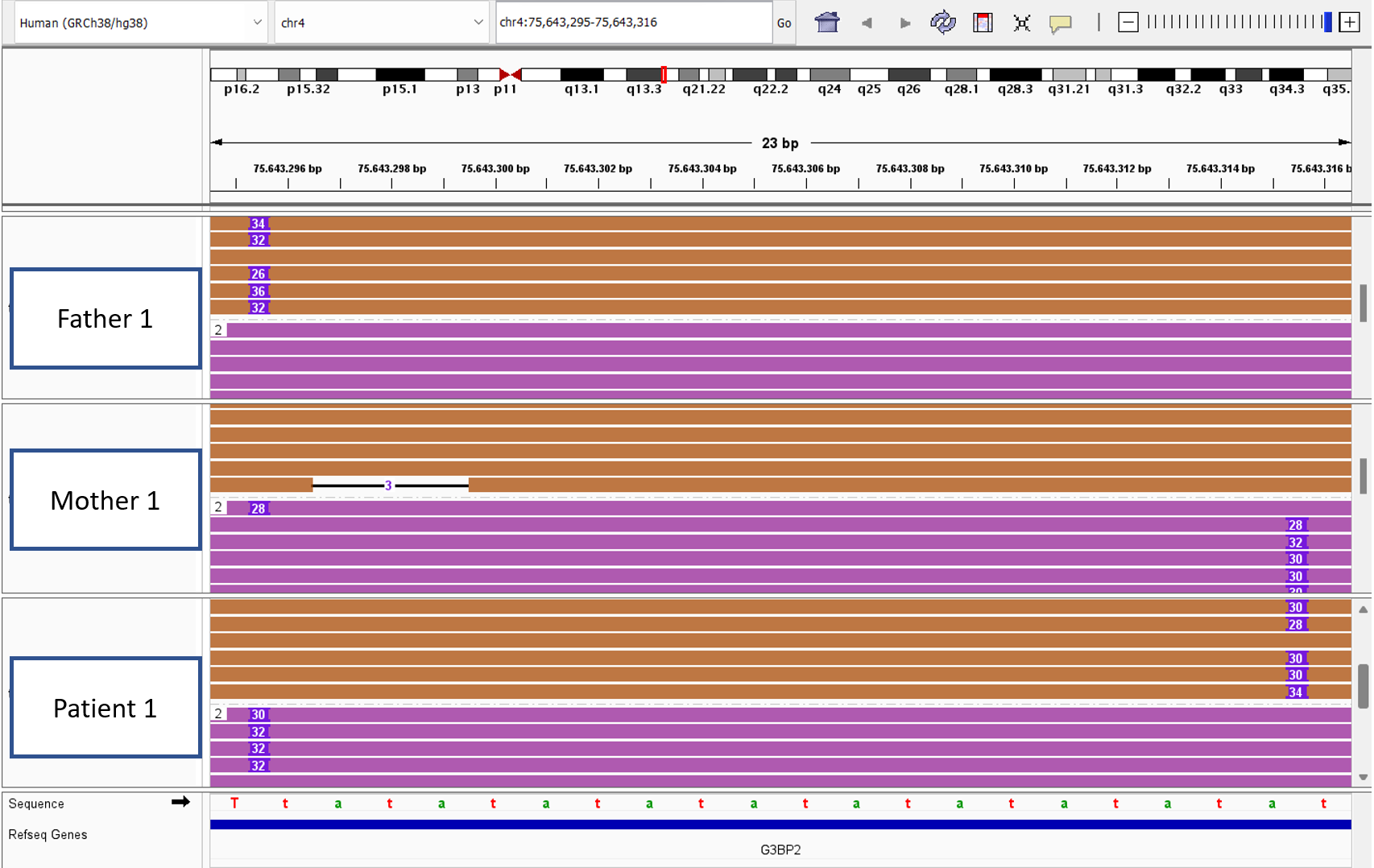
**Supplementary figure 4: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by Straglr on chromosome X:6,532,979.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by Straglr. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with. Reads that are indicated in grey could not be haplotagged.



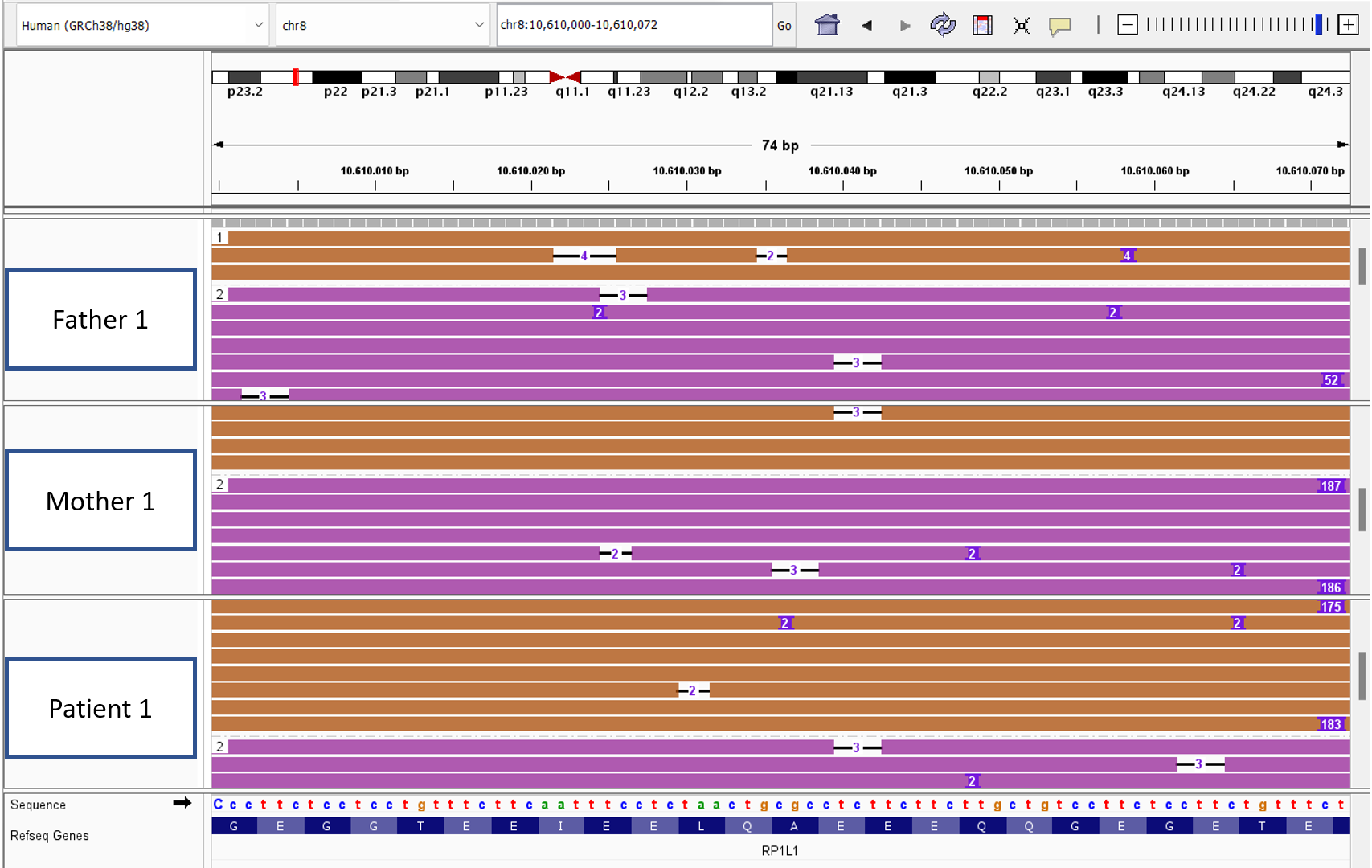
**Supplementary figure 5: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by Straglr on chromosome X:8,169,563.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by Straglr. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with. Reads that are indicated in grey could not be haplotagged.



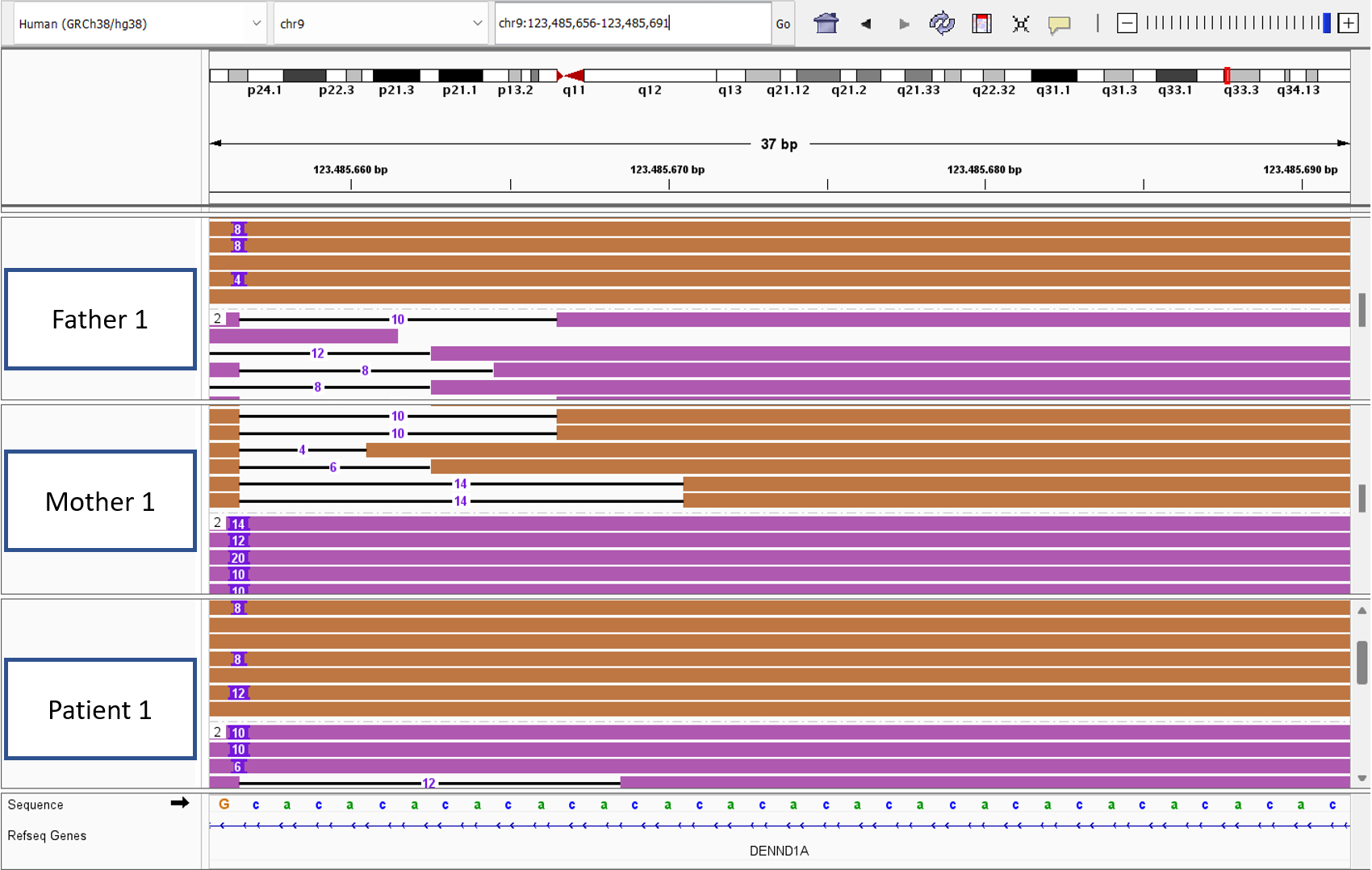
**Supplementary figure 6: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 2.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



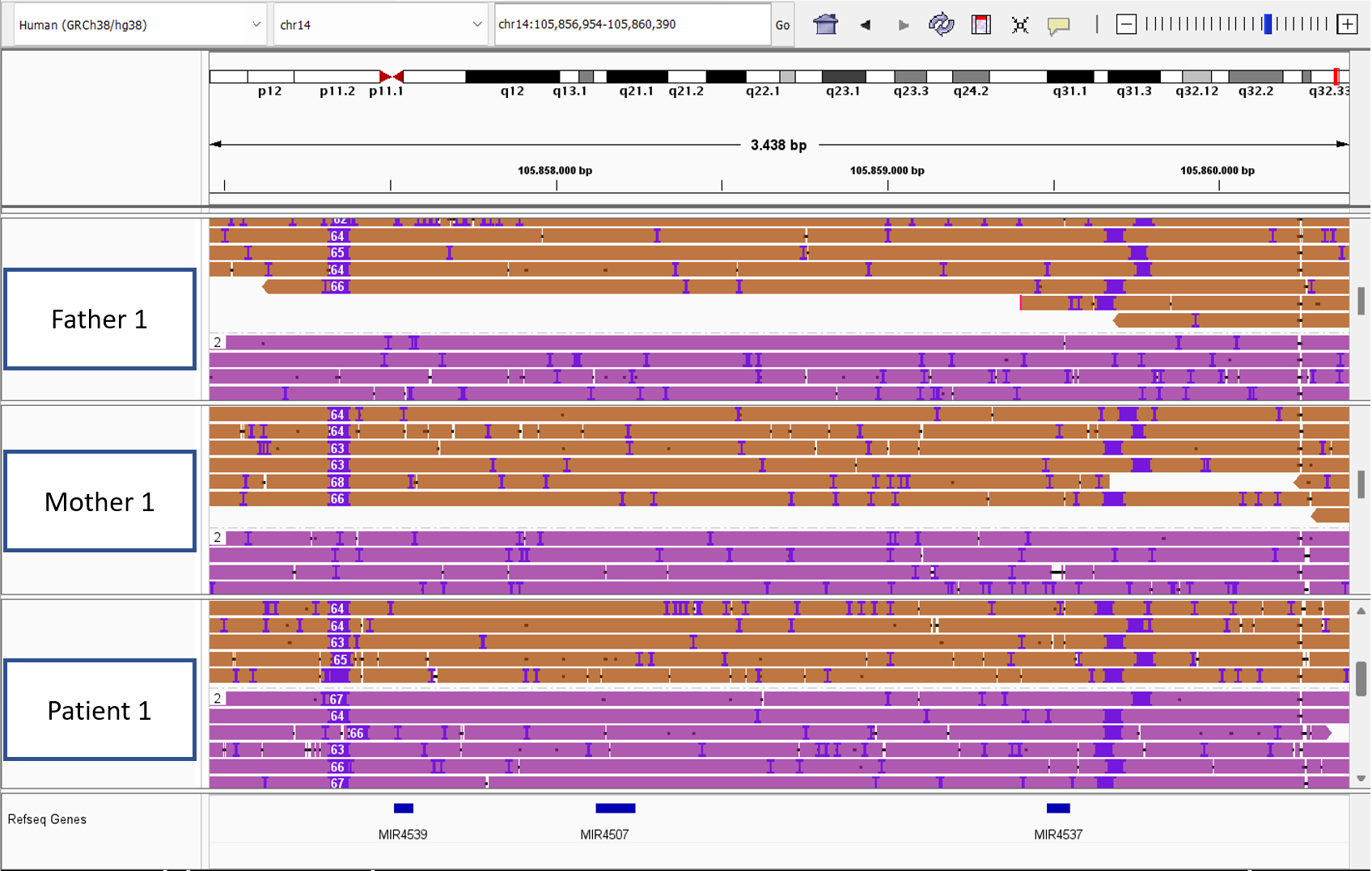
**Supplementary figure 7: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 4.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



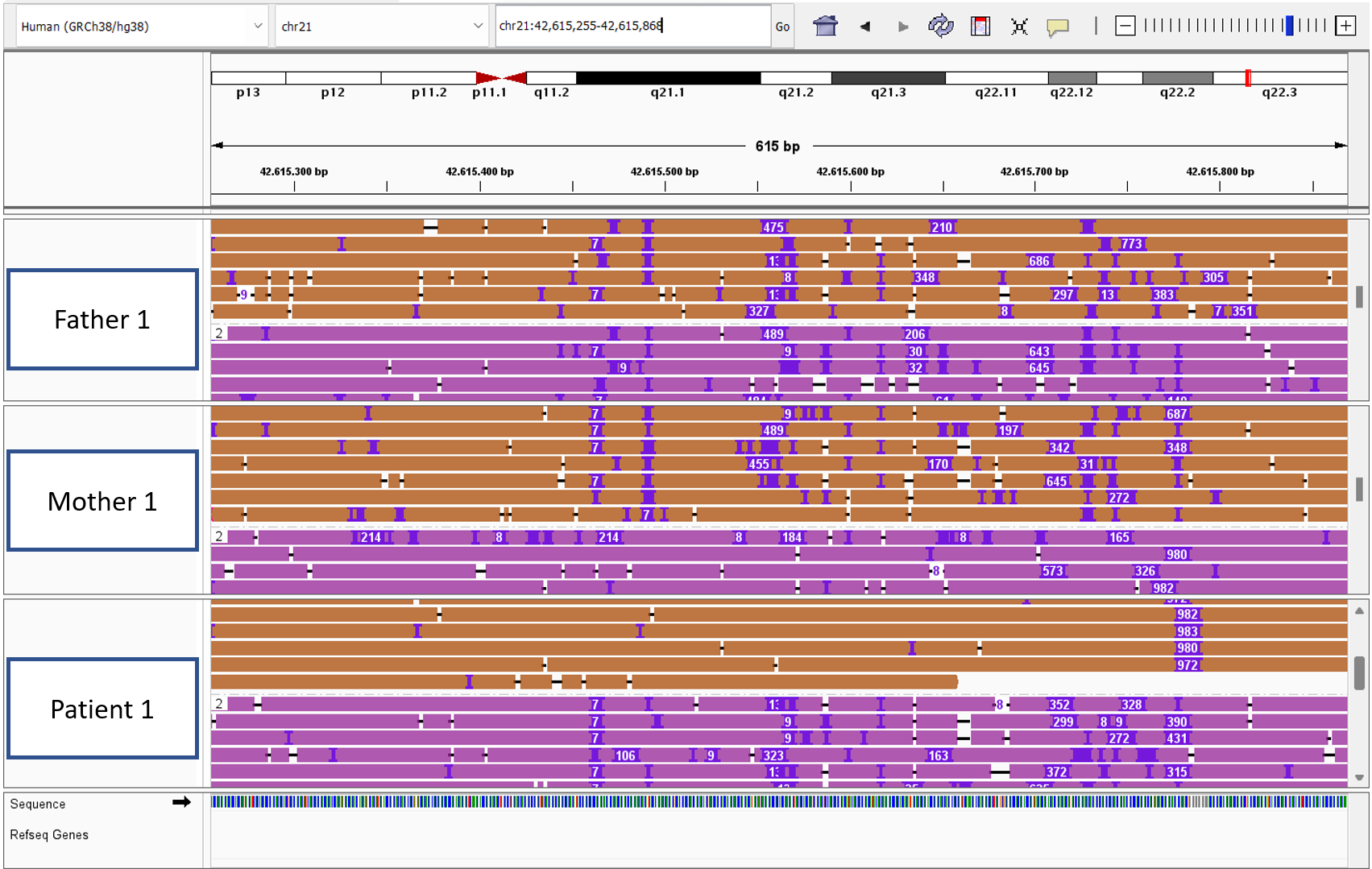
**Supplementary figure 8: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 8.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



**Supplementary figure 9: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 9.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



**Supplementary figure 10: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 14.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.



**Supplementary figure 11: Sequencing data Patient 1 and her parents in the region of the tandem-repeat, that was identified by tandem-genotypes on chromosome 21.** This screenshot from IGV shows Oxford Nanopore Sequencing reads that were mapped in the genomic region containing the tandem-repeat associated that was detected in the genome-wide analysis by tandem-genotypes. The blue rectangles with numbers inside represent insertions of the respective number of base pairs. The first, second and third track respectively show reads from the mother and father of the first patient, and the first patient (trio 19.22059). The genomic location is shown in the top right rectangle, and the sequence of the reference genome is depicted at the bottom of the figure. The narrow rectangles show the insertion sequence of one random insertion from each track. The colors of the reads indicate the haplotype they are associated with.

# Supplementary tables

**Supplementary table 1: Statistical metrics for detected lengths of known tandem-repeats in all analyzed individuals.** The first column shows the disease that the tandem-repeat is associated with. The second column shows the standard deviation of the detected tandem-repeats in repeated units, while the third column shows the average detected length of the tandem-repeat in repeated units.

|  |  |  |
| --- | --- | --- |
| Disease | Standard deviation | mean |
| SPD1 | 0.72 | 14.77 |
| SCA17 | 2.13 | 36.9 |
| GDPAG | 3.55 | 12.72 |
| SCA2 | 0.95 | 22.45 |
| SCA6 | 1.61 | 11.9 |
| FXS | 3.76 | 29.4 |
| HPE5 | 1.3 | 15.35 |
| BSS | 4.22 | 83.83 |
| SCA12 | 2.32 | 12.03 |
| SCA37 | 33.43 | 5.05 |
| FRDA | 3.48 | 10.83 |
| SCA8 | 3.41 | 13.65 |
| EPM1 | 1.34 | 3.62 |
| SCA1 | 1.93 | 30.74 |
| SCA31 | 337.19 | 113.14 |
| SCA3 | 5.33 | 20.09 |
| SCA36 | 2.27 | 8.4 |
| HFG | 0.99 | 13.55 |
| FRAXE | 2.63 | 51.44 |
| SCA7 | 2.84 | 11.71 |

**Supplementary table 2: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 19.22059 in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 19.22059 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 30.8 | 30.7 | 30.7 | 29/30 | 30/31 | 30/31 | 29 |
| SCA2 | 21.7 | 21.8 | 21.9 | 22/23 | 22/23 | 22/23 | 23 |
| SCA3 | 19.0 | 17.0 | 20.4 | 20/22 | 14/20 | 20/22 | 14 |
| SCA6 | 13.2 | 13.2 | 13.2 | 13/14 | 13/15 | 13/14 | 13 |
| SCA7 | 10.3 | 10.2 | 12.4 | 10/11 | 10/11 | 10/13 | 10 |
| SCA8 | 10.9 | 12/18 | 11.7 | 12/14 | 13/18 | 8/13 | 15 |
| SCA12 | 10.6 | 10.5 | 10.7 | 11/12 | 11/12 | 11/12 | 10 |
| SCA17 | 37.4 | 37.2 | 36.2 | 38/40 | 37/38 | 36/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 4/6 | 4/5 | 0/6 | 0 |
| SCA36 | 7.9 | 8.2 | 8.3 | 9/10 | 10/11 | 10/11 | 8 |
| SCA37 | n.d | n.d. | n.d. | -2/-1 | -2/-1 | -7/-2 | 0 |
| FXS | 29.4 | 30.3 | 28.9 | 30/33 | 30/33 | 29/31 | 20 |
| FRAXE | 51.0 | 52.8 | 49.7 | 51/55 | 52/55 | 50/51 | 49 |
| BSS | n.d. | n.d. | n.d. | 80/84 | 80/83 | 82/84 | 5 |
| FRDA | 12.2 | 11.9 | 12.2 | 8/10 | 9/10 | 8/9 | 6 |
| SPD1 | 15.4 | 15.6 | 15.4 | 14/15 | 14/15 | 14/15 | 14 |
| GDPAG | 14.0 | 14.8 | 13.3 | 15/16 | 15/16 | 8/15 | 16 |
| HFG | n.d | n.d. | n.d. | 12/14 | 13/14 | 13/14 | 14 |
| HPE5 | 14.7 | 14.9 | 14.6 | 15/18 | 15/16 | 15/16 | 15 |
| EPM1 | 3.8 | 3.8 | 4.7 | 2/3 | 2/3 | 2/3 | 3 |

**Supplementary table 3: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 15.00728 in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 15.00728as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 30.6 | 31.2 | 30.5 | 30/31 | 30/31 | 29/30 | 29 |
| SCA2 | 22.1 | 21.6 | 21.7 | 22/23 | 22/23 | 22/24 | 23 |
| SCA3 | 13.6 | 13.6 | 15.4 | 8/14 | 9/14 | 9/16 | 14 |
| SCA6 | 11.6 | 11.7 | 12.1 | 11/12 | 12/13 | 11/12 | 13 |
| SCA7 | 11.4 | 11.4 | 11/19 | 10/12 | 10/12 | 9/12 | 10 |
| SCA8 | 15.9 | 16.7 | 14.2 | 15/16 | 15/16 | 13/15 | 15 |
| SCA12 | 10/16 | 10.6 | 11/16 | 11/16 | 10/11 | 11/16 | 10 |
| SCA17 | 32/38 | 37.7 | 32/37 | 32/38 | 38/39 | 32/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 6/432 | 1/6 | 5/424 | 0 |
| SCA36 | 7.5 | 7.5 | 7.6 | 7/10 | 7/10 | 7/10 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -2/1 | -2/-1 | 1/2 | 0 |
| FXS | 28.4 | 28/37 | 30.0 | 30/31 | 31/38 | 30/31 | 20 |
| FRAXE | 48.9 | 49.7 | 52.6 | 49/51 | 50/52 | 51/53 | 49 |
| BSS | n.d. | n.d. | n.d. | 80/84 | 83/84 | 82/86 | 5 |
| FRDA | 9/14 | 11.0 | 8.7 | 8/9 | 9/10 | 7/8 | 6 |
| SPD1 | 15.4 | 15.5 | 15.7 | 14/15 | 14/15 | 14/15 | 14 |
| GDPAG | 12.4 | 14.3 | 11/16 | 8/15 | 14/15 | 8/17 | 16 |
| HFG | n.d. | n.d. | n.d. | 11/14 | 13/14 | 13/14 | 14 |
| HPE5 | 7.4 | 14.7 | 14.7 | 15/16 | 15/16 | 15/17 | 15 |
| EPM1 | 5.7 | 5.5 | 5.0 | 3/3 | 3/3 | 4/5 | 3 |

**Supplementary table 4: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 19.06383** **in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 19.06383 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 29.1 | 29.9 | 29.6 | 29/30 | 27/33 | 29/30 | 29 |
| SCA2 | 21.9 | 21.8 | 21.8 | 22/23 | 23/24 | 22/24 | 23 |
| SCA3 | 23.3 | 20/25 | 21.7 | 23/26 | 21/26 | 23/26 | 14 |
| SCA6 | 12.3 | 11.2 | 11.4 | 11/14 | 11/12 | 11/13 | 13 |
| SCA7 | 11/20 | 11/19 | 10.9 | 10/13 | 10/13 | 10/11 | 10 |
| SCA8 | 16.0 | 17.9 | 14.8 | 16/17 | 17/18 | 13/16 | 15 |
| SCA12 | 12.3 | 10.7 | 14/19 | 11/13 | 10/11 | 13/20 | 10 |
| SCA17 | 37.4 | 38.5 | 36.9 | 37/38 | 37/38 | 37/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 4/10 | -2/4 | 5/10 | 0 |
| SCA36 | 7.3 | 7.3 | 8.1 | 7/8 | 7/10 | 8/11 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -7/-1 | -7/-7 | -1/1 | 0 |
| FXS | 31.4 | 30.2 | 27.4 | 32/33 | 30/32 | 28/29 | 20 |
| FRAXE | 49.2 | 50.7 | n.d. | 49/50 | 48/55 | 57/59 | 49 |
| BSS | n.d. | n.d. | n.d. | 79/83 | 80/82 | 81/90 | 5 |
| FRDA | 11.0 | 12.4 | 9.1 | 9/10 | 8/10 | 7/10 | 6 |
| SPD1 | 15.8 | 15.5 | 15.1 | 14/15 | 14/15 | 14/15 | 14 |
| GDPAG | 9/16 | 10.1 | 12.4 | 6/16 | 6/8 | 9/14 | 16 |
| HFG | n.d. | n.d. | n.d. | 13/15 | 13/15 | 13/14 | 14 |
| HPE5 | 15.2 | 14.9 | 15.7 | 15/16 | 15/17 | 16/17 | 15 |
| EPM1 | 5.7 | 4.6 | 5.5 | 3/3 | 2/3 | 3/3 | 3 |

**Supplementary table 5: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 17.10085 in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 17.10085as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 30.6 | 30.8 | 30.4 | 30/31 | 30/33 | 28/31 | 29 |
| SCA2 | 21.5 | 21.9 | 21.8 | 22/25 | 22/26 | 22/23 | 23 |
| SCA3 | 20/29 | 21.8 | 23/29 | 21/30 | 21/23 | 23/30 | 14 |
| SCA6 | 9.3 | 9.4 | 9.0 | 7/11 | 7/11 | 7/12 | 13 |
| SCA7 | 10/18 | 10.0 | 10/18 | 10/11 | 10/11 | 10/11 | 10 |
| SCA8 | 13.9 | 12.1 | 17.4 | 13/14 | 12/13 | 14/17 | 15 |
| SCA12 | 11/16 | 13.5 | 10.5 | 10/15 | 10/15 | 10/11 | 10 |
| SCA17 | 36.1 | 36.2 | 35.9 | 35/37 | 36/37 | 26/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 7/976 | 1161/1591 | 7/8 | 0 |
| SCA36 | 6.5 | 6.4 | 6.9 | 7/8 | 7/8 | -1/0 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -8/2 | 2/211 | -8/-7 | 0 |
| FXS | 29.8 | 29.8 | 30.4 | 30/31 | 30/32 | 31/34 | 20 |
| FRAXE | 50.0 | 50/58 | 47.9 | 49/52 | 51/58 | 47/50 | 49 |
| BSS | n.d. | n.d. | n.d. | 84/93 | 82/85 | 83/95 | 5 |
| FRDA | 11.8 | 13/22 | 10.5 | 9/10 | 9/17 | 7/9 | 6 |
| SPD1 | 15.5 | 15.5 | 15.6 | 13/14 | 14/15 | 14/15 | 14 |
| GDPAG | 11/18 | 17.5 | 11.3 | 8/13 | 16/19 | 8/9 | 16 |
| HFG | n.d. | n.d. | n.d. | 14/15 | 13/16 | 13/14 | 14 |
| HPE5 | 15.4 | 15.3 | 14.6 | 15/17 | 15/16 | 15/17 | 15 |
| EPM1 | 3.9 | 4.9 | 4.7 | 2/3 | 2/3 | 2/3 | 3 |

**Supplementary table 6: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 13.13599 in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 13.13599 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 37.0 | 31/37 | 30.0 | 29/36 | 31/36 | 29/30 | 29 |
| SCA2 | 21.5 | 21.7 | 21.9 | 22/23 | 22/24 | 22/23 | 23 |
| SCA3 | 14/23 | 21.9 | 15/21 | 14/23 | 23/24 | 14/22 | 14 |
| SCA6 | 12.6 | 11.8 | 11.6 | 12/14 | 12/13 | 11/12 | 13 |
| SCA7 | 9.7 | 10.9 | 10.1 | 9/10 | 10/11 | 10/11 | 10 |
| SCA8 | 14.4 | 17.9 | 16.3 | 13/15 | 13/17 | 8/15 | 15 |
| SCA12 | 10.6 | 10.7 | 10.7 | 10/11 | 10/11 | 10/11 | 10 |
| SCA17 | 37.0 | 36.7 | 37.5 | 37/38 | 37/41 | 38/39 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 4/5 | -2/5 | 3/4 | 0 |
| SCA36 | 6.8 | 7.7 | 7/11 | 7/10 | 7/10 | 7/12 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -3/7 | -3/8 | -7/7 | 0 |
| FXS | 28.1 | 29.0 | 19.3 | 30/31 | 30/31 | 20/21 | 20 |
| FRAXE | 50.1 | 49.7 | 48.0 | 50/53 | 50/55 | 48/49 | 49 |
| BSS | n.d. | n.d. | n.d. | 76/81 | 80/83 | 78/85 | 5 |
| FRDA | 11.9 | 11.9 | 10.9 | 8/10 | 8/9 | 8/9 | 6 |
| SPD1 | 15.3 | 15.4 | 15.4 | 14/15 | 14/15 | 14/15 | 14 |
| GDPAG | 14.8 | 14.4 | 11/17 | 15/17 | 14/15 | 8/18 | 16 |
| HFG | n.d. | n.d. | n.d. | 13/14 | 12/14 | 12/14 | 14 |
| HPE5 | 14.7 | 14.6 | 14.9 | 15/16 | 15/16 | 15/16 | 15 |
| EPM1 | 5.6 | 5.5 | 5.0 | 3/3 | 3/3 | 2/3 | 3 |

**Supplementary table 7: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 20.03510 in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 20.03510 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 31.9 | 31.0 | 32.1 | 32/33 | 29/32 | 31/34 | 29 |
| SCA2 | 22.1 | 21.4 | 21.3 | 22/23 | 22/23 | 22/23 | 23 |
| SCA3 | 14/25 | 13/26 | 14/26 | 14/27 | 14/27 | 14/27 | 14 |
| SCA6 | 12.1 | 11.0 | 12.1 | 12/14 | 11/12 | 12/13 | 13 |
| SCA7 | 10/17 | 10/19 | 12.1 | 9/10 | 9/10 | 10/13 | 10 |
| SCA8 | 6/13 | 14/20 | 6.5 | 5/13 | 13/19 | 5/8 | 15 |
| SCA12 | 10.6 | 11.8 | 10.0 | 10/11 | 10/13 | 10/11 | 10 |
| SCA17 | 36.5 | 35.8 | 36.7 | 37/38 | 36/37 | 37/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 4/6 | 4/10 | 4/6 | 0 |
| SCA36 | 9.1 | 9.0 | 9.7 | 10/12 | 10/12 | 11/12 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -2/16 | -3/-2 | 8/16 | 0 |
| FXS | 23/29 | 29.0 | 22.2 | 24/30 | 29/31 | 23/24 | 20 |
| FRAXE | 52.0 | 51.4 | 52.2 | 52/55 | 52/56 | 52/53 | 49 |
| BSS | n.d. | n.d. | n.d. | 82/85 | 82/84 | 82/85 | 5 |
| FRDA | 16.1 | 9/20 | 10/15 | 10/17 | 7/19 | 9/10 | 6 |
| SPD1 | 15.2 | 15.9 | 15.5 | 14/16 | 14/15 | 14/15 | 14 |
| GDPAG | 10.9 | 15.9 | 13.0 | 8/9 | 8/16 | 8/15 | 16 |
| HFG | n.d. | n.d. | n.d. | 13/14 | 13/14 | 13/15 | 14 |
| HPE5 | 14.8 | 14.6 | 15.4 | 15/16 | 15/16 | 15/16 | 15 |
| EPM1 | 4.6 | 4.5 | 5.1 | 2/3 | 2/3 | 2/3 | 3 |

**Supplementary table 8: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 18.00975** **in reads mapped to human genome 38 (hg38).** This table shows the reported tandem-repeat lengths for all individuals in trio 18.00975 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in hg38 |
| SCA1 | 29.4 | 30.5 | 29.8 | 29/30 | 29/31 | 29/30 | 29 |
| SCA2 | 21.4 | 21.6 | 23.0 | 22/23 | 22/23 | 22/25 | 23 |
| SCA3 | 15/21 | 15/21 | 21/27 | 14/21 | 14/21 | 21/27 | 14 |
| SCA6 | 12.0 | 11.9 | 12.0 | 12/15 | 12/13 | 12/13 | 13 |
| SCA7 | 10.9 | 11/18 | 12/20 | 10/12 | 10/12 | 10/13 | 10 |
| SCA8 | 11.0 | 9/16 | 13.7 | 8/13 | 8/15 | 13/16 | 15 |
| SCA12 | 14.6 | 11.9 | 11/16 | 14/15 | 10/13 | 10/15 | 10 |
| SCA17 | 37.1 | 37.7 | 37.2 | 37/38 | 38/39 | 37/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 0/4 | 5/8 | 0/2 | 0 |
| SCA36 | 7.8 | 7.3 | 7.4 | 8/9 | 8/10 | 8/10 | 8 |
| SCA37 | n.d. | n.d. | n.d. | -3/4 | -3/1 | 1/4 | 0 |
| FXS | 33.7 | 22/33 | 30.2 | 24/34 | 23/34 | 31/32 | 20 |
| FRAXE | 50.8 | 51.3 | 47.9 | 51/53 | 51/53 | 49/50 | 49 |
| BSS | n.d. | n.d. | n.d. | 84/90 | 83/91 | 85/96 | 5 |
| FRDA | 10.6 | 9/14 | 12/22 | 7/10 | 8/9 | 10/18 | 6 |
| SPD1 | 15.3 | 15.2 | 15.4 | 14/14 | 14/15 | 13/14 | 14 |
| GDPAG | 11.8 | 12.1 | 11/18 | 8/14 | 8/11 | 9/19 | 16 |
| HFG | n.d. | n.d. | n.d. | 13/15 | 12/14 | 13/14 | 14 |
| HPE5 | 14.9 | 15.0 | 14.8 | 15/18 | 15/16 | 15/16 | 15 |
| EPM1 | 4/6 | 5.5 | 4/7 | 2/3 | 3/3 | 2/2 | 3 |

**Supplementary table 9: Detected tandem-repeat lengths by tandem-genotypes and Straglr in trio 19.22059** **in reads mapped to the telomere-to-telomere reference genome (T2T).** This table shows the reported tandem-repeat lengths for all individuals in trio 19.22059 as detected by tandem-genotypes and Straglr. The second to fourth columns depict Straglr’s results, the fifth to seventh columns depict the reported TR lengths by tandem-genotypes (“tg”), and the last column depicts the number of repeats in the reference. The symbol “/” delineates values from heterozygous calls and if a TR could not be detected by a tool, this is indicated with “n.d.”.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Proband  Straglr | Mother Straglr | Father Straglr | Proband tg | Mother tg | Father tg | Repeats in T2T |
| SCA1 | 30.8 | 30.7 | 30.6 | 29/31 | 29/30 | 29/30 | 29 |
| SCA2 | 21.9 | 21.8 | 21.9 | 22/24 | 22/23 | 22/24 | 22 |
| SCA3 | 5.9 | 15/5 | 6.4 | 10/14 | 4/14 | 8/14 | 14 |
| SCA6 | 13.2 | 13.2 | 13.2 | 13/14 | 13/15 | 13/14 | 13 |
| SCA7 | 10.3 | 18/10 | 12.4 | 10/11 | 10/11 | 10/13 | 10 |
| SCA8 | 10.9 | 18/12 | 11.7 | 9/14 | 14/19 | 9/14 | 9 |
| SCA12 | 10.6 | 10.8 | 10.8 | 6/10 | 6/10 | 6/10 | 10 |
| SCA17 | 37.4 | 37.2 | 36.2 | 38/40 | 36/38 | 36/38 | 38 |
| SCA31 | n.d. | n.d. | n.d. | 0/2 | -1/0 | -4/1 | 0 |
| SCA36 | 7.2 | 7.5 | 7.4 | 7/8 | 7/8 | 7/8 | 7 |
| SCA37 | n.d. | n.d. | n.d. | 6/7 | 6/7 | 1/6 | 0 |
| FXS | 29.4 | 30.1 | 29.3 | 30/31 | 30/33 | 30/31 | 30 |
| FRAXE | 51.0 | 52.4 | 49.7 | 50/55 | 51/54 | 50/51 | 50 |
| BSS | 33.2 | 32.5 | 33.9 | 31/33 | 31/34 | 33/34 | 34 |
| FRDA | 12.2 | 11.9 | 12.2 | 7/9 | 8/9 | 8/9 | 9 |
| SPD1 | 15.4 | 15.6 | 15.4 | 14/16 | 14/16 | 14/15 | 14 |
| GDPAG | 14.0 | 14.8 | 13.7 | 14/15 | 15/16 | 8/15 | 14 |
| HFG | n.d. | n.d. | n.d. | 11/12 | 11/12 | 11/12 | 12 |
| HPE5 | 14.5 | 14.9 | 14.6 | 15/16 | 15/16 | 15/16 | 15 |
| EPM1 | 3.8 | 6/4 | 4.7 | 2/3 | 2/3 | 2/3 | 3 |