

Supporting Material and Methods

Detailed analysis of *HTT* repeat elements in human blood using targeted amplification-free long-read sequencing

Ida Höijer, Yu-Chih Tsai, Tyson A. Clark, Paul Kotturi, Niklas Dahl, Eva-Lena Stattin, Marie-Louise Bondeson, Lars Feuk, Ulf Gyllensten and Adam Ameur.

Supporting Figures

Figure S1. Target designs. No-amp targeted sequencing target designs for the *ATXN10*, *FMR1* and *C9orf72* repeat regions.

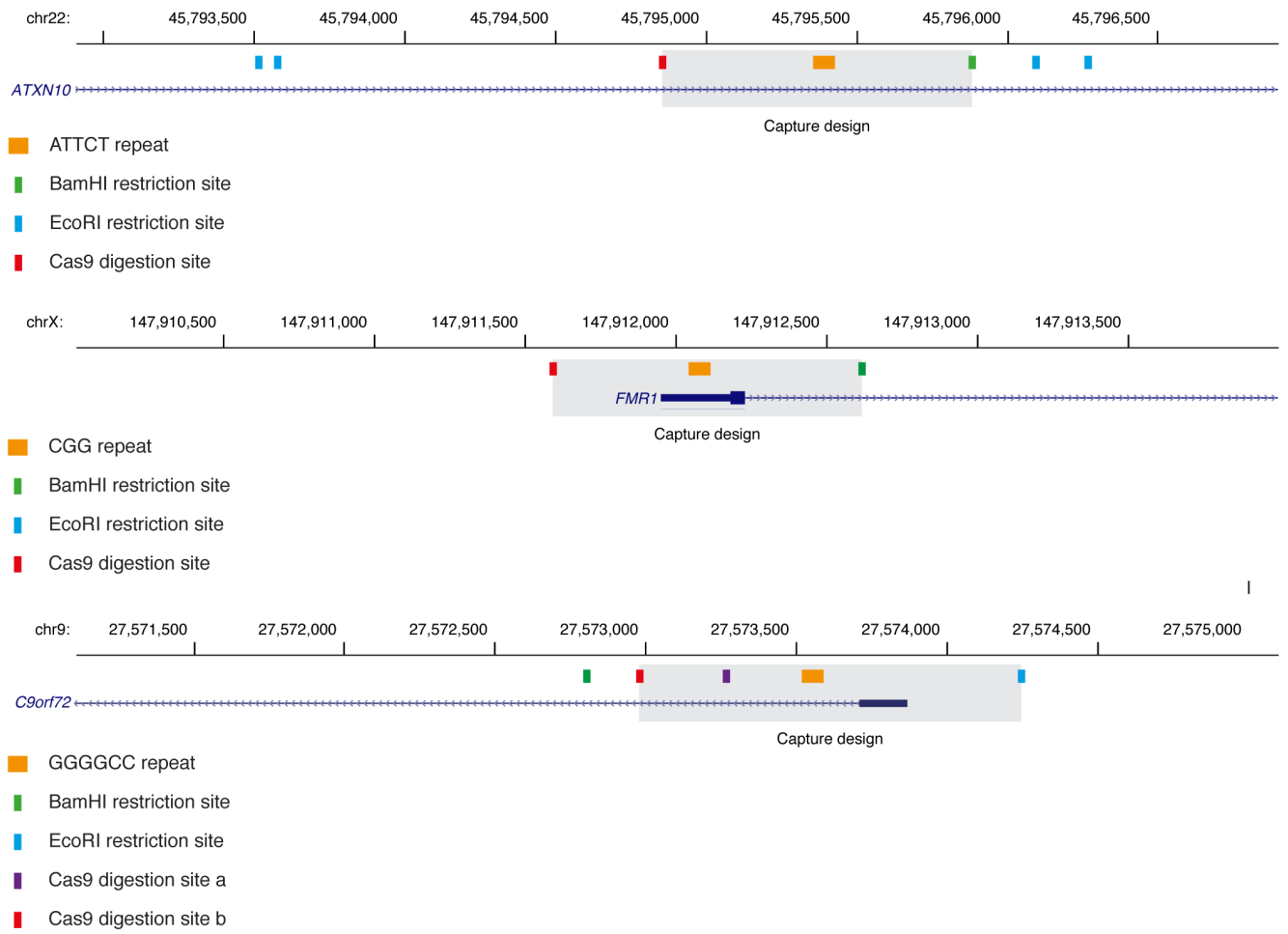


Figure S2. Coverage plots for replicates. Genome wide coverage plots for the HEK 293 replicates.

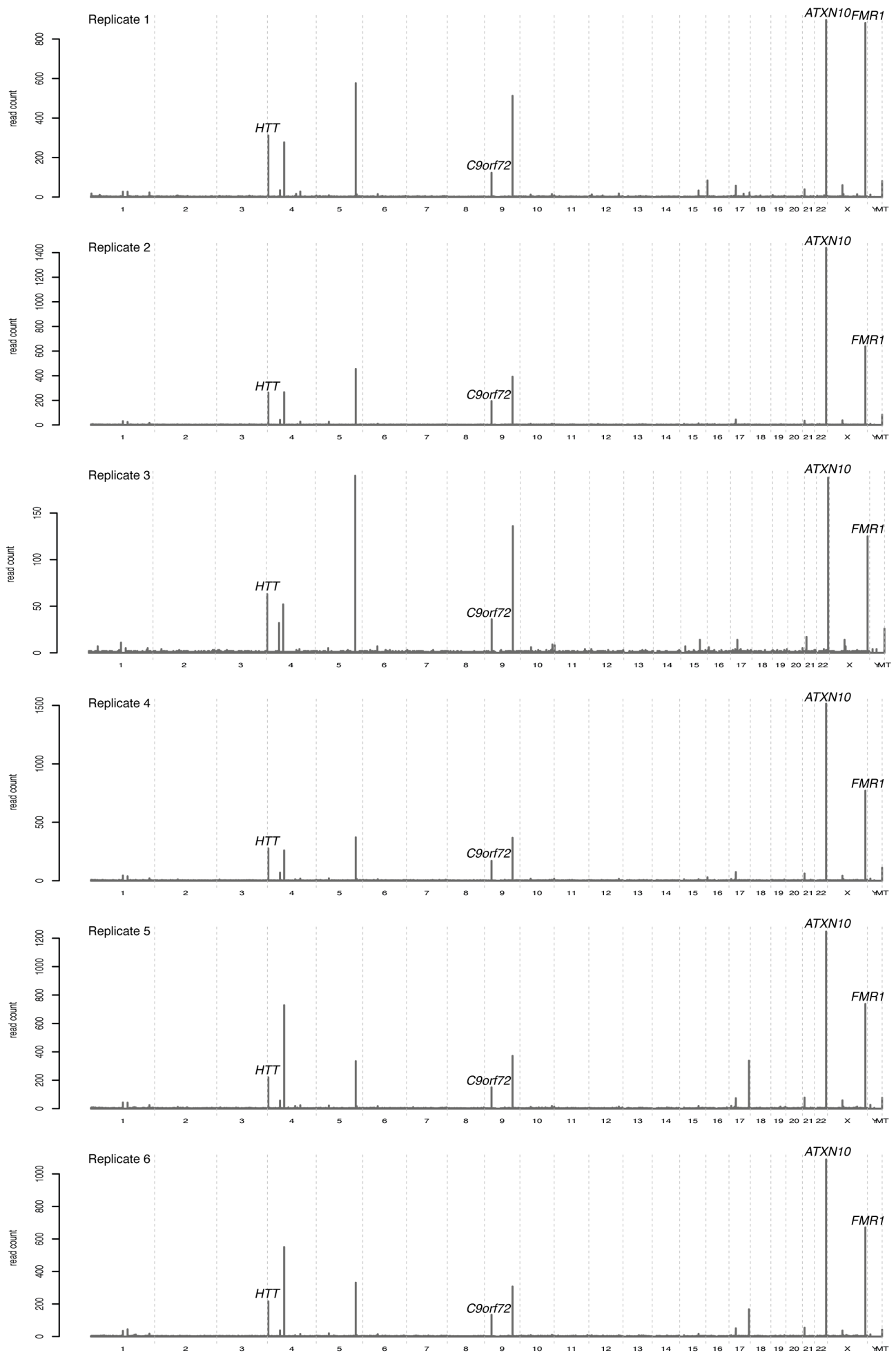


Figure S3. Off-target effects. Off-target effects caused by gRNA homology to other loci in the genome. A) Off-target effect caused by homology to the *HTT* gRNA in chromosome 5. **B)** Off-target effect in chromosome 9. Only genomes with SNP rs7861875 show the off-target effect caused by homology to the *ATXN10* gRNA. **C)** Off-target effect in chromosome 4 caused by homology to the *ATXN10* gRNA. Genomic coordinates refer to the GRCh38 reference genome.

A

```
chr5:154,409,392-154,409,411  AGAAGGCCTAAACTCACGGT
                               ||..|||||.|||||||
HTT gRNA                       AGCGGGCCCAAACUCACGGU
```

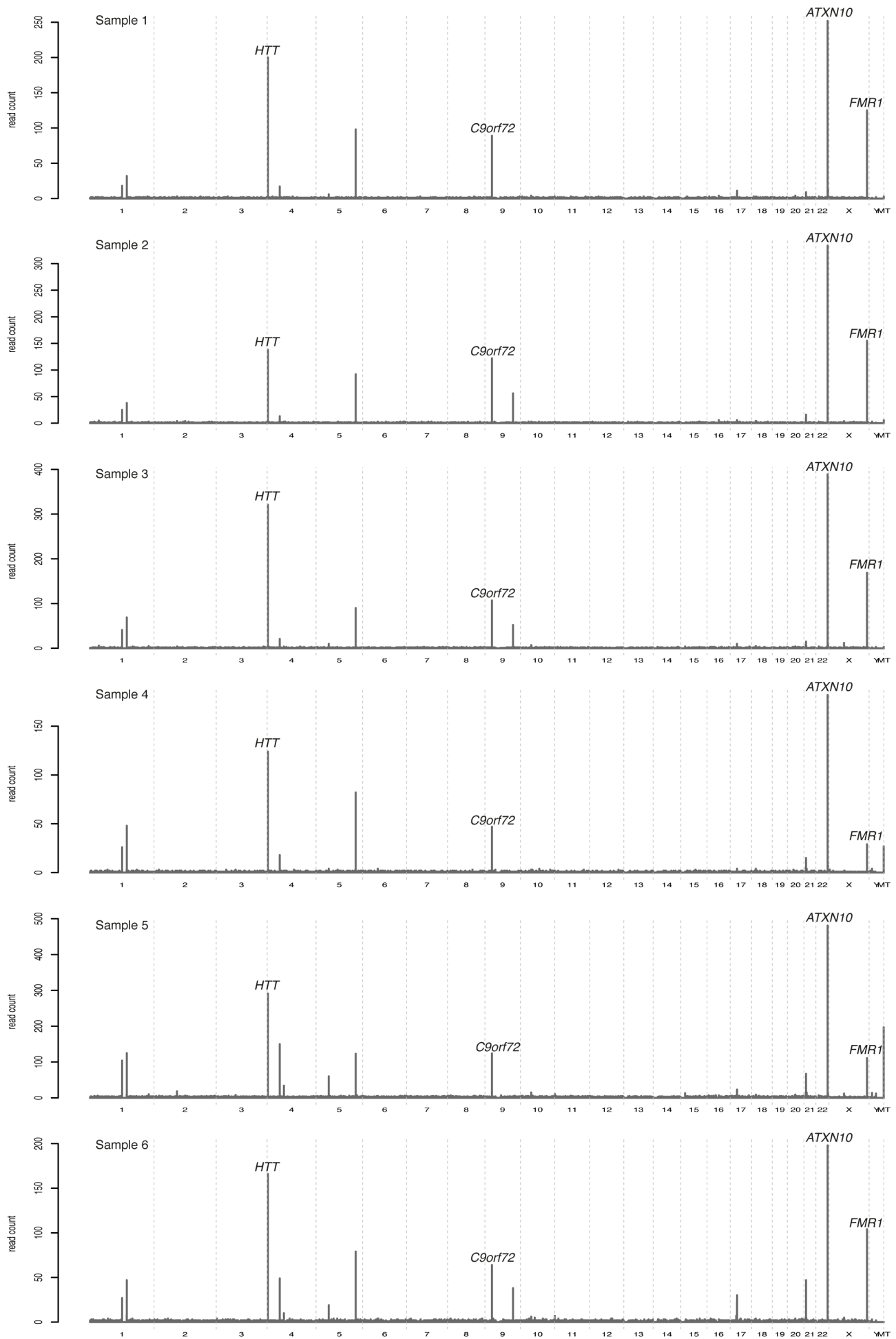
B

```
                               rs7861875
                               G
chr9:109,570,949-109,570,968  ATACAACATATCAGAATCCC
                               |||||. . . |||||
ATXN10 gRNA                    AUACAAAGGAUCAGAAUCCC
```

C

```
chr4:65,163,126-65,163,145  TCACAAAGGATAAGAATCCC
                               ..|||||||. . |||||
ATXN10 gRNA                    AUACAAAGGAUCAGAAUCCC
```

Figure S4. Coverage plots for HD samples. Genome wide coverage plots for the 11 HD samples.



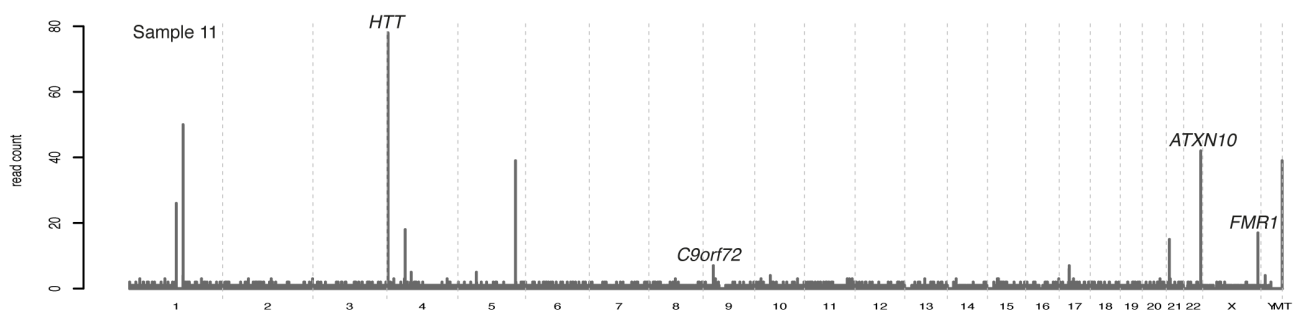
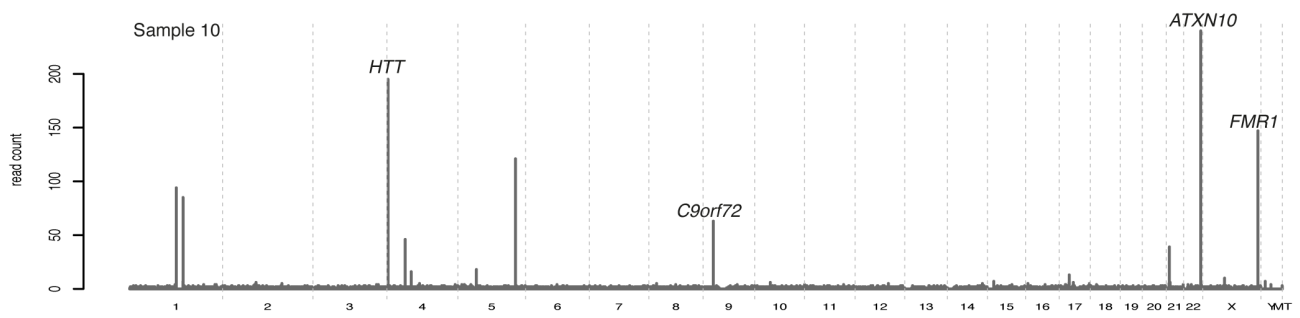
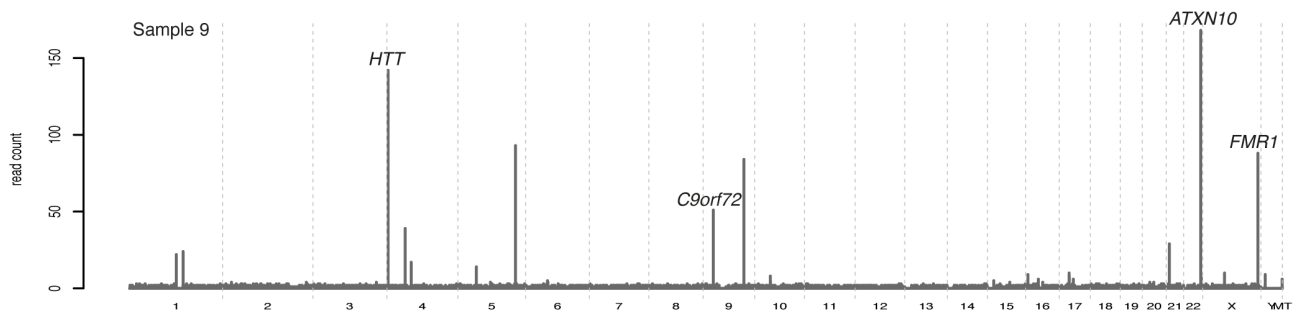
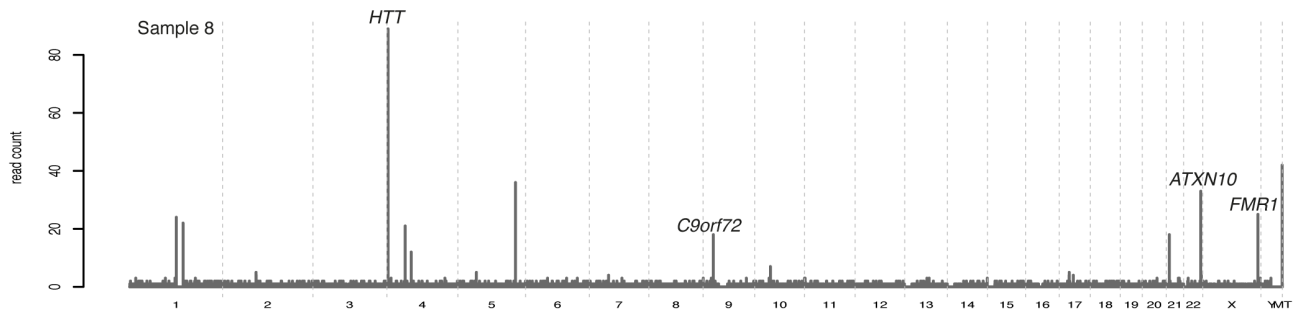
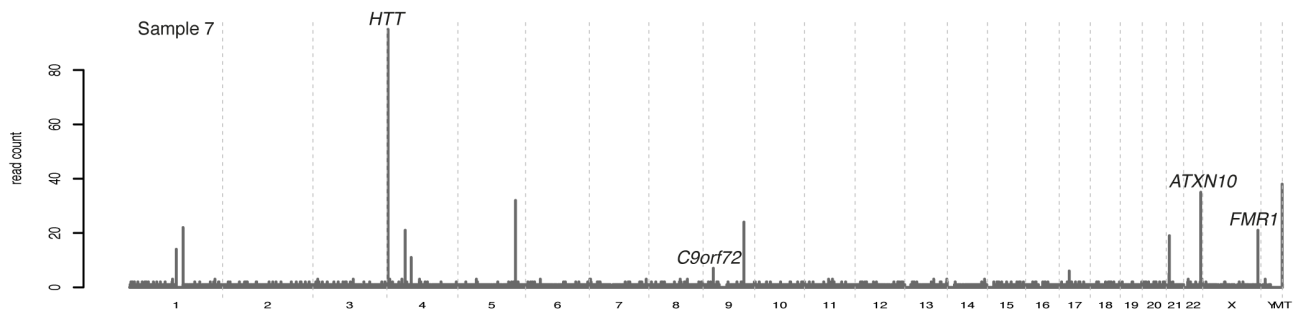
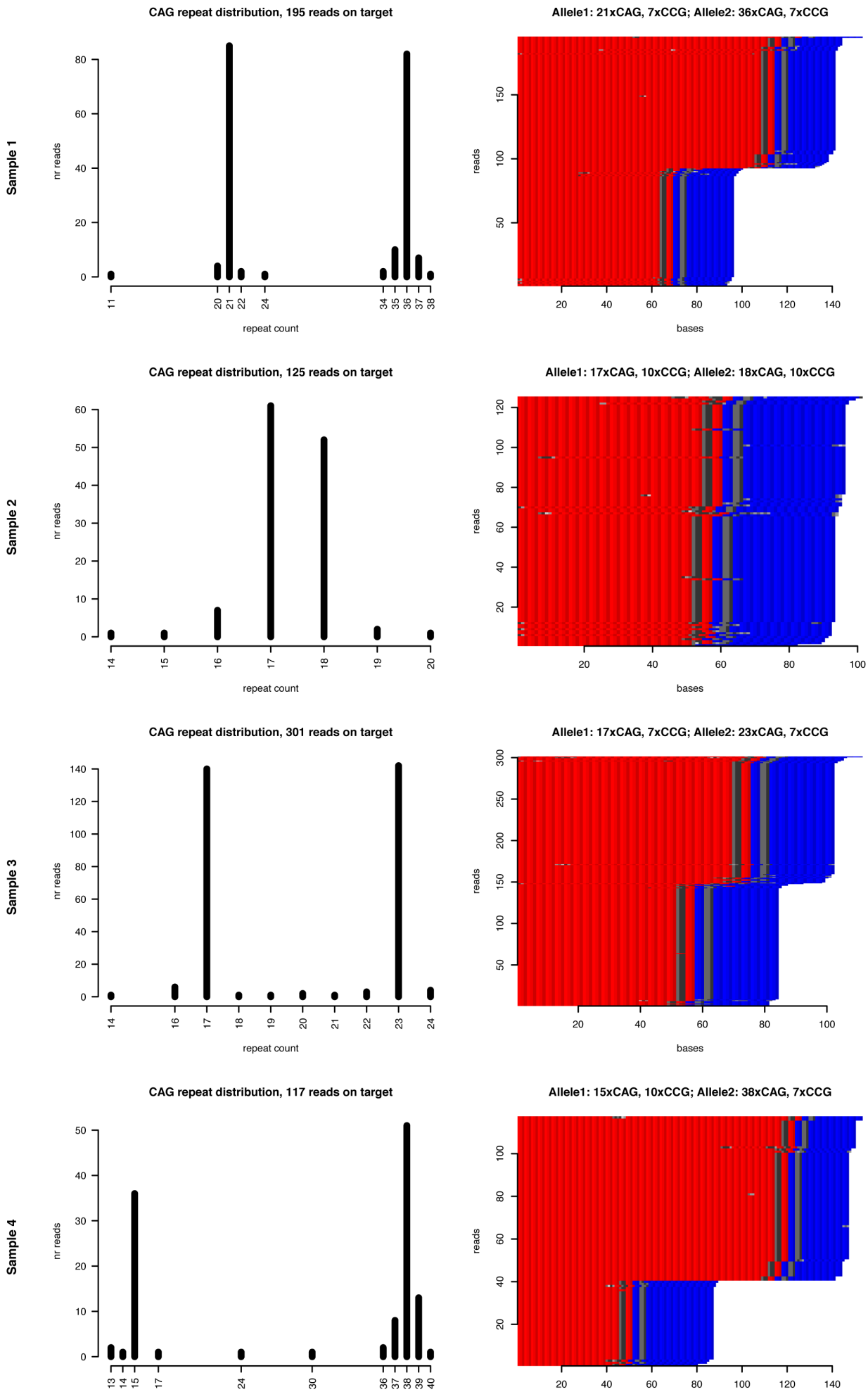
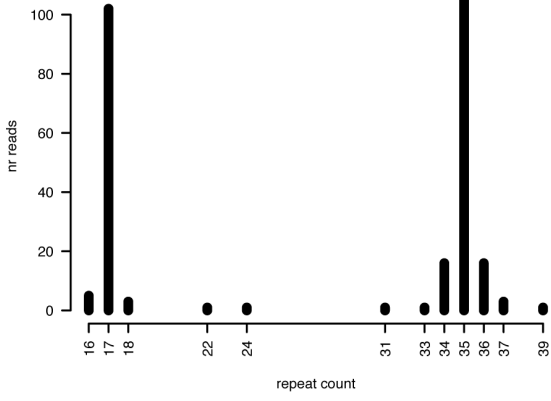


Figure S5. Allelic distribution of CAG repeats and repeat-content plots for 11 HD samples. The histograms in the left column show the allelic CAG distribution for all HD samples. The repeat-content plots in the right column show CAG repeats in red and CCG repeats in blue. Sequences that are neither CAG nor CCG are shown in grey.

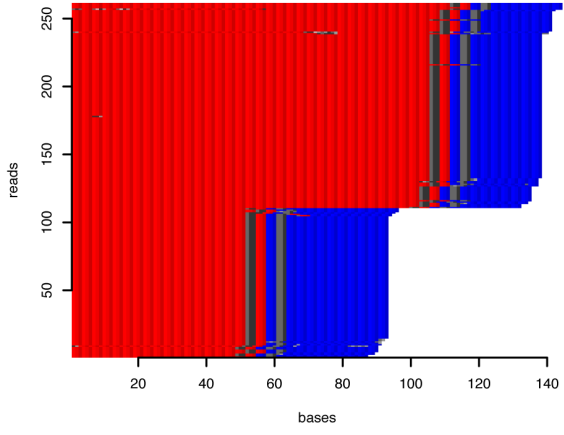


Sample 5

CAG repeat distribution, 261 reads on target

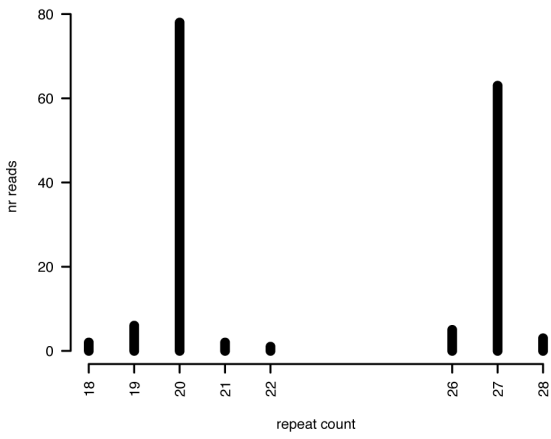


Allele1: 17xCAG, 10xCCG; Allele2: 35xCAG, 7xCCG

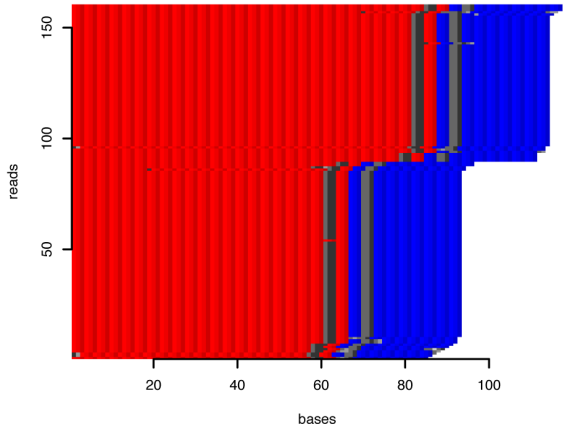


Sample 6

CAG repeat distribution, 160 reads on target

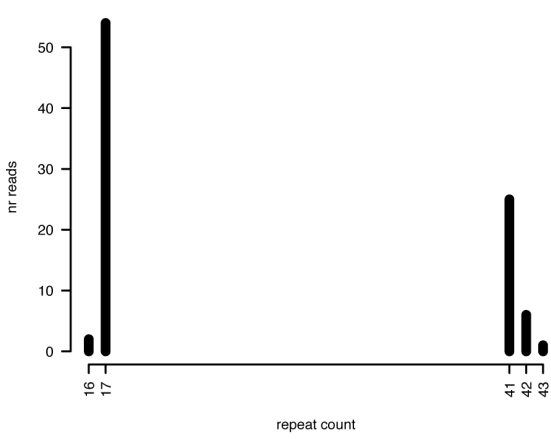


Allele1: 20xCAG, 7xCCG; Allele2: 27xCAG, 7xCCG

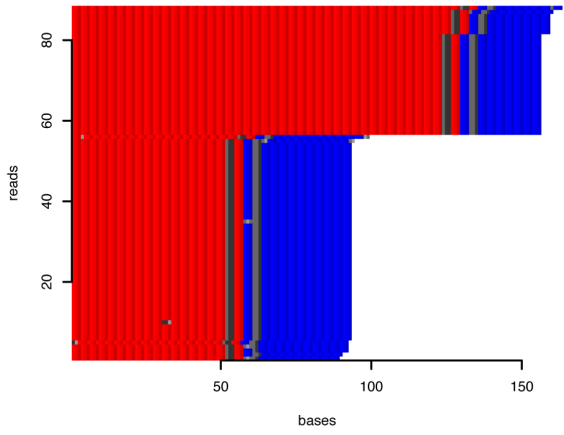


Sample 7

CAG repeat distribution, 88 reads on target

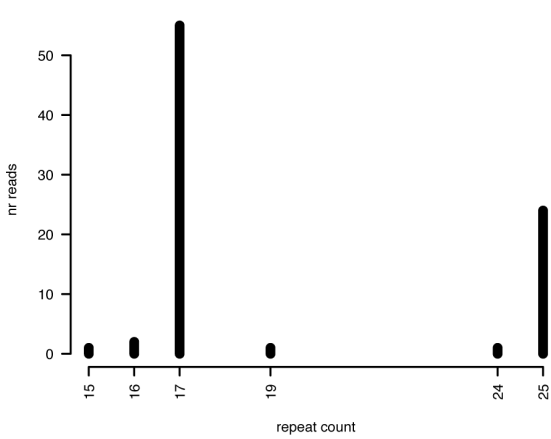


Allele1: 17xCAG, 10xCCG; Allele2: 41xCAG, 7xCCG

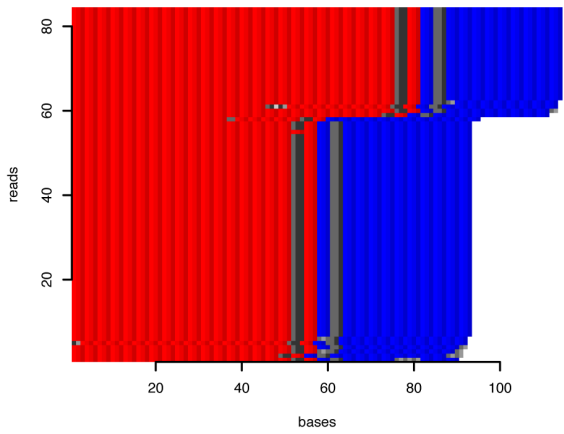


Sample 8

CAG repeat distribution, 84 reads on target

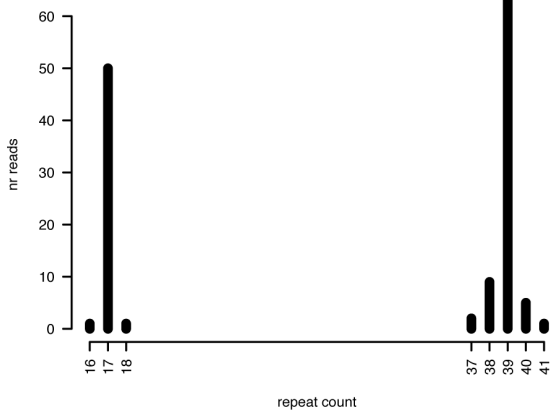


Allele1: 17xCAG, 10xCCG; Allele2: 25xCAG, 9xCCG

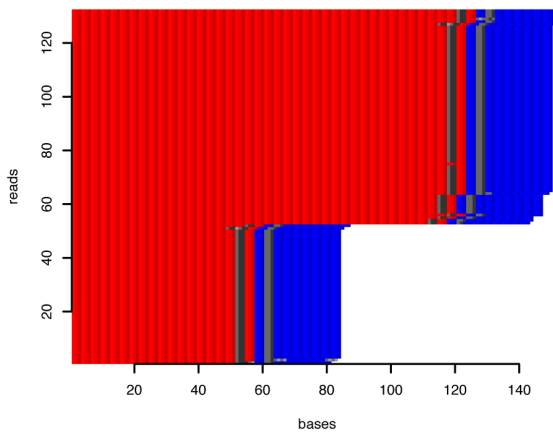


Sample 9

CAG repeat distribution, 132 reads on target

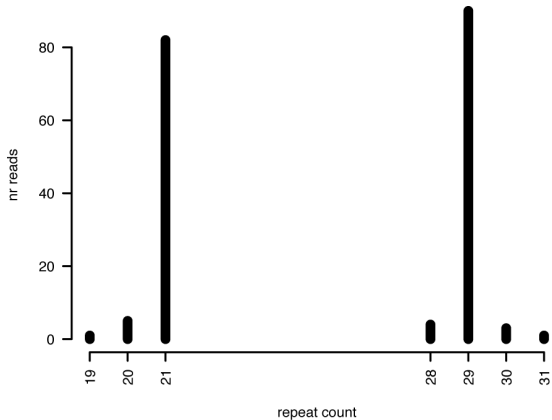


Allele1: 17xCAG, 7xCCG; Allele2: 39xCAG, 7xCCG

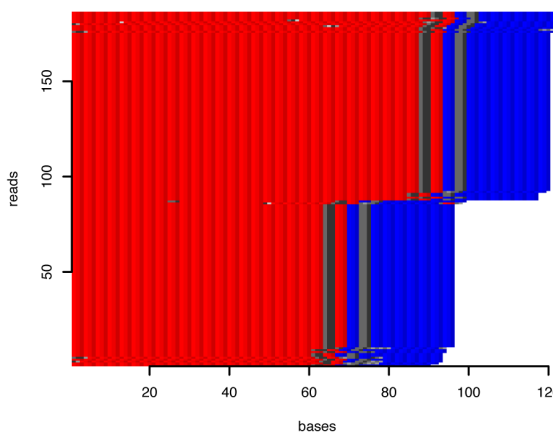


Sample 10

CAG repeat distribution, 186 reads on target

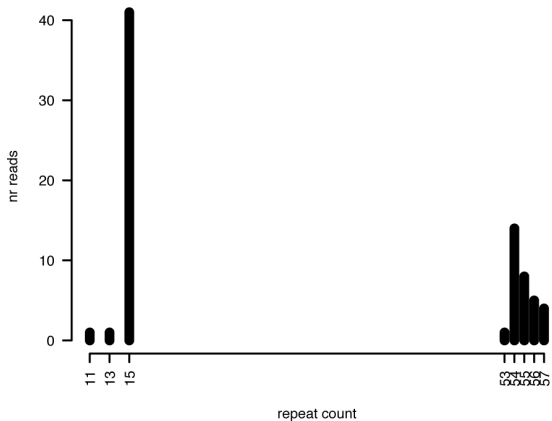


Allele1: 21xCAG, 7xCCG; Allele2: 29xCAG, 7xCCG



Sample 11

CAG repeat distribution, 75 reads on target



Allele1: 15xCAG, 10xCCG; Allele2: 54xCAG, 7xCCG

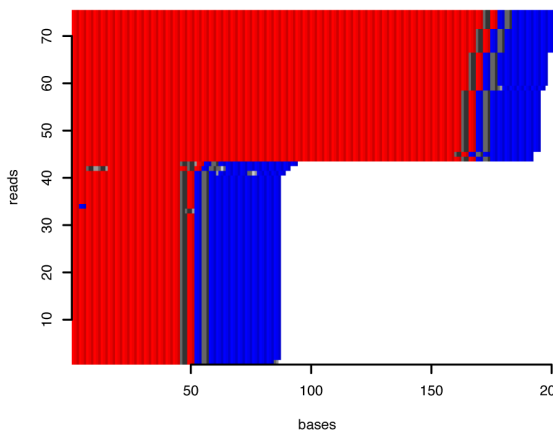
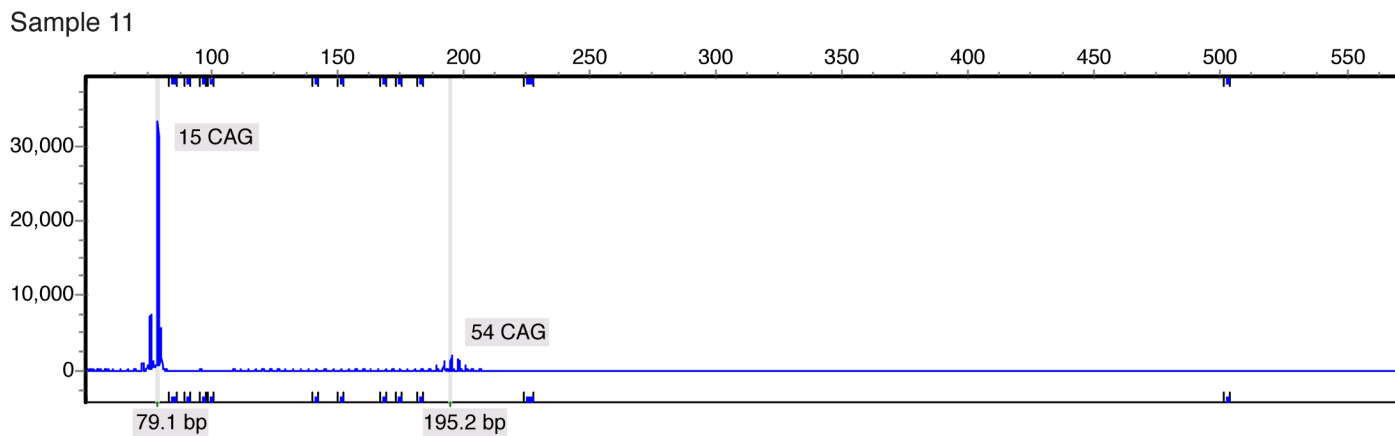


Figure S6. Fragment analysis results for sample 11. Fragment analysis result for sample 11 showing two alleles with 15 and 54 CAG repeats respectively.



Supporting Tables

Table S1. Sample library preparation information.

Sample	Input DNA (μ g)	Amount into capture (ng)	Complexity reduction restriction enzymes	gRNAs
Replicate 1	20	764	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Replicate 2	20	743	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Replicate 3	20	398	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Replicate 4	20	588	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Replicate 5	20	668	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Replicate 6	20	651	KpnI, EcoRV, MfeI, SpeI	<i>ATXN10, FMR1, HTT, C9orf72a</i>
Sample 1	5.5	570	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 2	4.9	468	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 3	5.1	378	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 4	3.8	258	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 5	5.7	792	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 6	5.7	588	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 7	5.9	714	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 8	7.7	762	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 9	5.7	582	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 10	5.9	666	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>
Sample 11	7.4	786	KpnI, EcoRV	<i>ATXN10, FMR1, HTT, C9orf72b</i>

Table S2. Reads on-target and total CCS reads for HEK 293 replicates.

Sample	CCS reads	Total CCS reads on-target	<i>C9orf72</i> (GGGGCC = 3)*			<i>ATXN10</i> (ATTCT = 14)*			<i>FMR1</i> (CGG = 19)*			<i>HTT</i> (CAG = 19)*		
			CCS reads on-target	(GGGGCC) _n Allele 1	(GGGGCC) _n Allele 2	CCS reads on-target	(ATTCT) _n Allele 1	(ATTCT) _n Allele 2	CCS reads on-target	(CGG) _n Allele 1	(CGG) _n Allele 2	CCS reads on-target	(CAG) _n Allele 1	(CAG) _n Allele 2
Replicate 1	31218	1401	59	2	4	580	16	16	480	28	29	282	17	17
Replicate 2	27408	1906	114	2	4	1143	16	16	400	28	29	249	17	17
Replicate 3	11194	324	26	2	4	147	16	16	90	28	29	61	17	17
Replicate 4	38879	1863	98	2	4	1079	16	16	426	28	29	260	17	17
Replicate 5	39863	1559	80	2	4	851	16	16	426	28	29	202	17	17
Replicate 6	29731	1386	76	2	4	727	16	16	383	28	29	200	17	17

* Number of repeats in human reference genome GRCh38

Table S3. Reads on-target and total CCS reads for 11 patient samples.

Sample	CCS reads	Total CCS reads on-target	<i>C9orf72</i> (GGGGCC = 3)*			<i>ATXN10</i> (ATTCT = 14)*			<i>FMR1</i> (CGG = 19)*			<i>HTT</i> (CAG = 19)*		
			CCS reads on-target	(GGGGCC) _n Allele 1	(GGGGCC) _n Allele 2	CCS reads on-target	(ATTCT) _n Allele 1	(ATTCT) _n Allele 2	CCS reads on-target	(CGG) _n Allele 1	(CGG) _n Allele 2	CCS reads on-target	(CAG) _n Allele 1	(CAG) _n Allele 2
Sample 1	5902	595	85	2	2	214	14	15	101	28	30	195	21	36
Sample 2	7317	644	114	2	2	295	13	14	110	27	28	125	17	18
Sample 3	9892	868	96	2	2	337	12	15	134	28	29	301	17	23
Sample 4	5775	333	43	7	15	155	14	17	18	34	35	117	15	38
Sample 5	33415	827	109	2	2	396	13	14	61	28	28	261	17	35
Sample 6	11590	452	60	2	8	164	14	17	68	22	28	160	20	27
Sample 7	4862	131	7	2	5	27	14	14	9	23	27	88	17	41
Sample 8	6025	137	16	2	2	25	14	14	12	35	37	84	17	25
Sample 9	10624	378	49	2	2	143	13	14	54	28	29	132	17	39
Sample 10	13500	548	60	2	2	202	14	16	100	28	35	186	21	29
Sample 11	5852	120	6	2	4	28	13	14	11	29	29	75	15	54

* Number of repeats in human reference genome GRCh38