The American Journal of Human Genetics, Volume *92*

Supplemental Data

Large *C9orf72* Hexanucleotide Repeat Expansions Are Seen in Multiple Neurodegenerative Syndromes and Are More Frequent Than Expected in the UK Population

Jon Beck, Mark Poulter, Davina Hensman, Jonathan D. Rohrer, Colin J. Mahoney, Gary Adamson, Tracy Campbell, James Uphill, Aaron Borg, Pietro Fratta, Richard W. Orrell, Andrea Malaspina, James Rowe, Jeremy Brown, John Hodges, Katie Sidle, James M. Polke, Henry Houlden, Jonathan M. Schott, Nick C. Fox, Martin N. Rossor, Sarah J. Tabrizi, Adrian M. Isaacs, John Hardy, Jason D. Warren, John Collinge, and Simon Mead

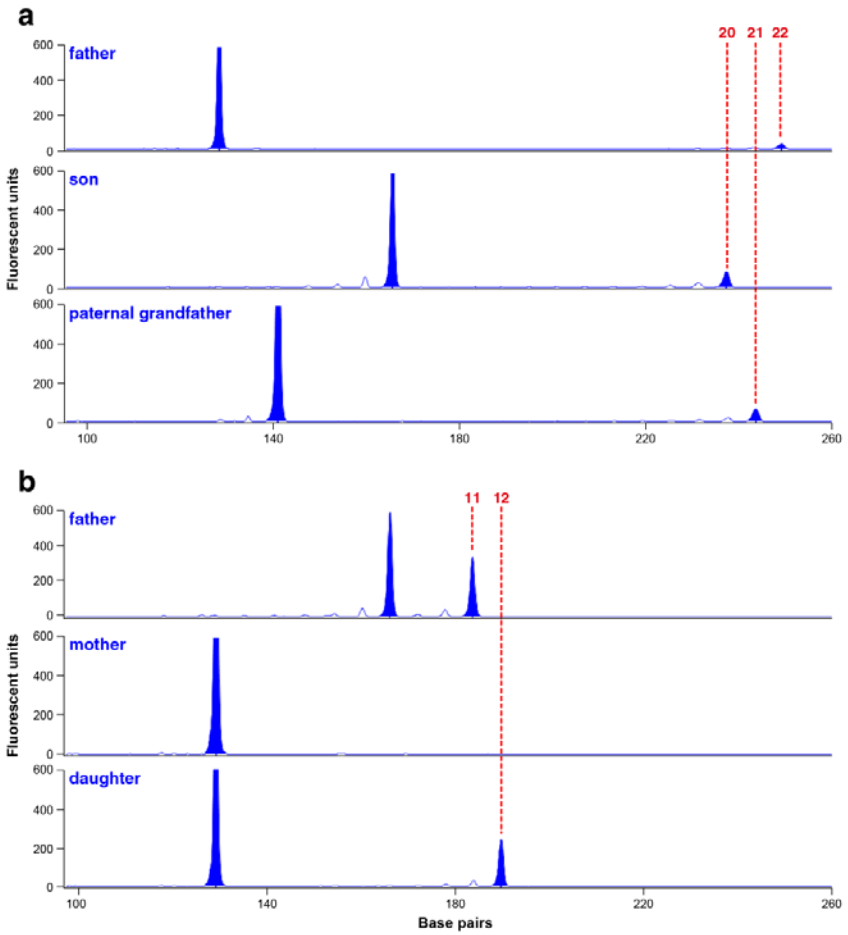

Figure S1. Sizing of Expansions in CEPH Families

Data from fluorescent labelled PCR followed by fragment length analysis on an ABI 3730xl automated sequencer from 2 CEPH families showing evidence of inter-generational repeat slippage. For clarity, the numbers of base pairs of alleles demonstrating slippage are also shown with repeat size in red text. A) CEPH family 1423 results showing slippage from paternal grandfather's 21 repeats up to father's 22 and then down to 20 repeats in his son. B) CEPH family 1420 showing slippage from father's 11 repeats to his daughter's 12 repeats.

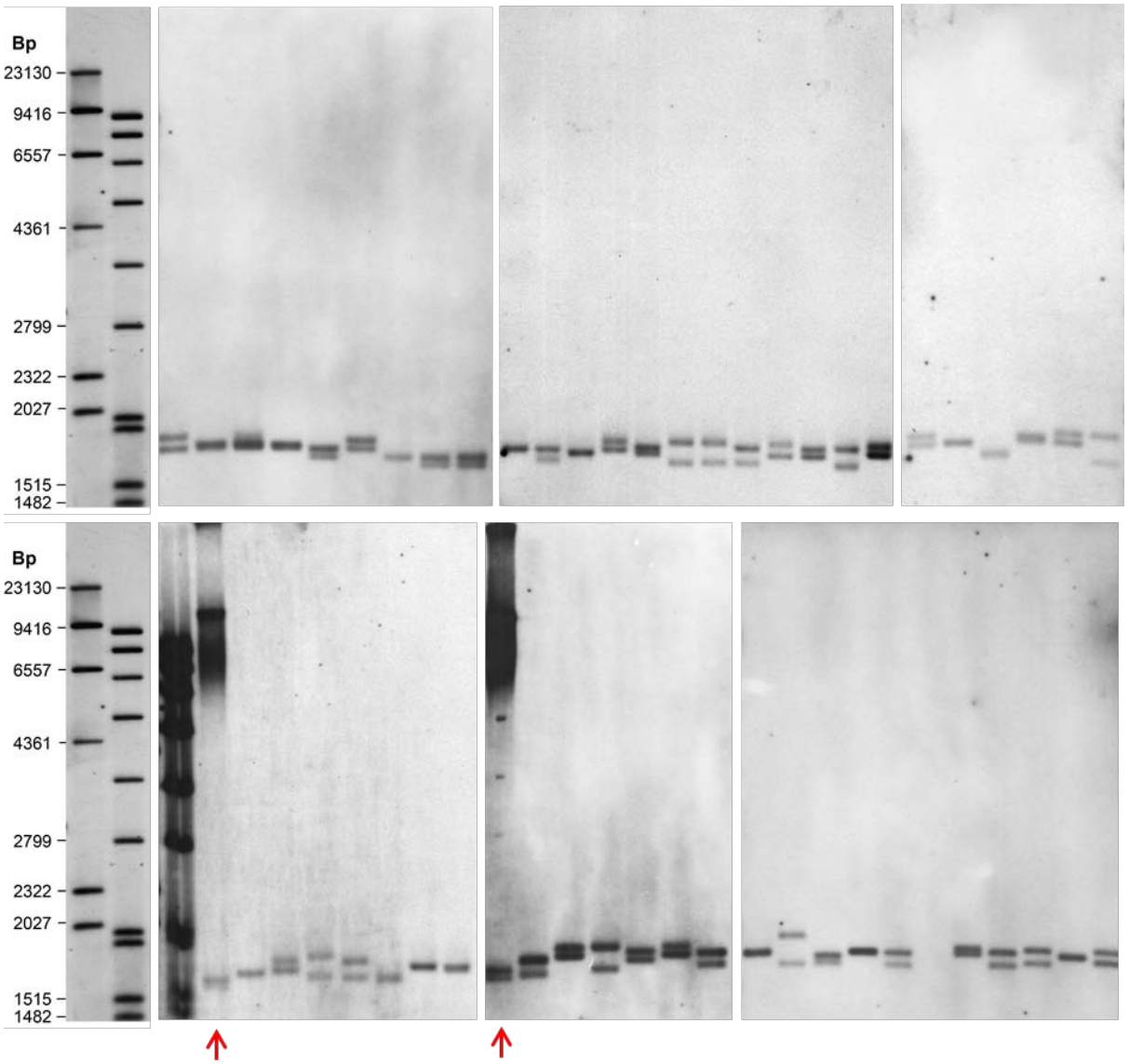

Figure S2. Control Southern Blotting of Samples in which rpPCR Detected Expansions < 32 Repeats

Southern blots for 50 rpPCR negative individuals shown with Ladder II and VII (Roche). Positive control individuals are indicated by arrows. The putative G4C2 containing repeat is the only signal seen in rpPCR negative controls. This signal is a useful internal control for blotting and hybridisation efficiency. The final block of controls contains a blank lane.

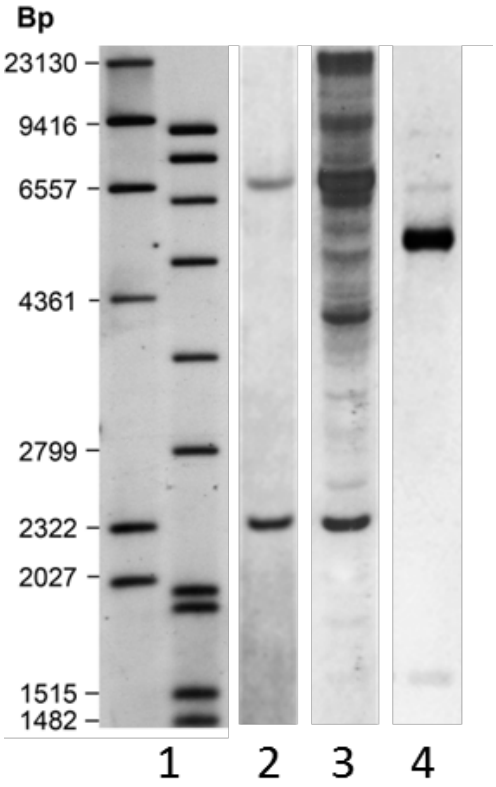

Figure S3. Southern Blots Highlighting the Sensitivity of the Modified Method

Blots to highlight the sensitivity and specificity of the $(GGGGCC)_5$ probe. Lymphoblastoid cell line DNA with expanded repeats was blotted and hybridised in various ways 1) Ladder II and Ladder VII (Roche), 2) Xba-1 digest with single copy probe, 3) Xba-1 digest with single copy probe and oligonucleotide probe combined and 4) Alu1/Dde1 double digest with oligonucleotide probe only. Sensitivity to the expanded allele (6.6Kb) is enhanced disproportionately when compared with the wild type allele (2.3kb) but specificity is lost (2 and 3). Sensitivity is maintained and specificity is regained by the use of an Alu1/Dde1 double digest (3 and 4). The sensitivity of the $(GGGGCC)_5$ probe is further highlighted by its ability to show pauciclonality represented by fainter bands of higher molecular weight than the main band.

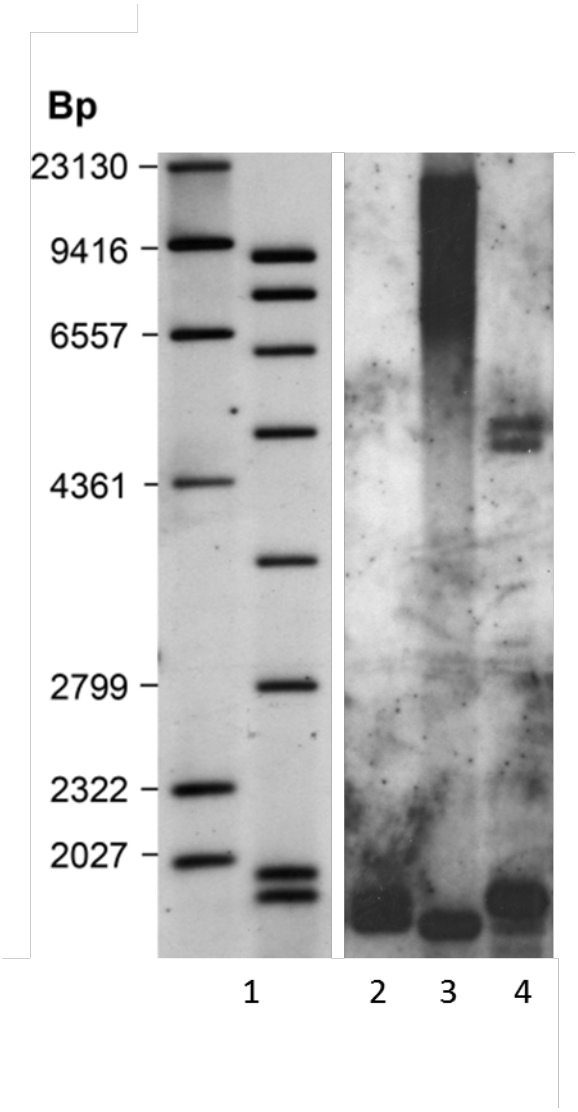

Figure S4. Southern Blot of Atypical rpPCR Positive Sample

Lane 1) Ladder II and VII (Roche), Lane 2) rpPCR negative control individual, Lane 3) rpPCR positive individual with typical smear pattern and Lane 4) rpPCR positive individual with two distinct bands.

Table S1. Data from Microsatellite Analysis in CEPH and Cases with *C9orf72* Expansions Showing a Lack of Association between the Risk Associated SNP rs3849942 and Any Microsatellite Marker

Marker name (D9S)	distance (Mb)	span (Mb)	Marker grouping (size bp)	CEPH rs3849942A	CEPH rs3849942G	P	CEPH rs3849942A/G heterozygotes	C9 expansion +ve rs3849942A/G heterozygotes	P	CEPH all rs3849942 genotypes	C9 expansion +ve all rs3849942 genotypes	P
1814	22.1	5.4	273, 275	12	51	0.92	28	45	0.19	76	64	0.25
			277	5	19		14	34		30	40	
			279, 283	2	11		10	9		18	14	
976	22.5	5	126, 128	10	43	0.82	24	47	0.62	60	54	0.45
			130, 132, 134	6	20		17	23		28	33	
			136, 138, 142	8	25		11	16		36	27	
171	24.5	3	160, 162, 164, 168	6	29	0.25	18	26	0.48	42	35	0.16
			170	9	32		18	30		45	43	
			172, 174, 176, 178	6	19		13	19		30	27	
			180, 182	1	0		1	7		1	7	
1121	25.4	2.1	170, 182	8	27	0.96	23	30	0.39	39	40	0.25
			194, 202	2	10		6	8		17	10	
			206	5	22		9	19		30	25	
			210, 214	4	17		10	27		28	39	
169	27.2	0.3	254, 258, 260, 262, 264	9	32	0.49	18	40	0.44	47	57	0.28
			266, 268, 270	4	17		11	11		23	15	
			272	6	9		11	22		27	27	
			274, 276, 278, 282	2	6		5	11		9	15	
C9ORF72 (27.5Mb)												
263	27.8	0.3	180	8	32	0.89	23	36	0.92	51	47	0.65
			182, 184	10	34		26	46		61	63	
			186, 188	2	10		3	6		12	8	
270	28.4	0.9	86, 88, 92	5	8	0.32	10	18	0.60	17	19	0.92
			94	8	34		24	35		53	54	
			96, 98, 100	8	31		15	33		48	46	
104	28.5	1	182, 184, 186, 188	3	30	0.07	25	37	0.86	51	47	0.41
			190, 192, 194	12	33		18	33		56	46	
			196, 198, 204	5	9		9	15		17	23	
147	31	3.5	181, 183, 185, 187	8	29	0.04	18	26	0.51	43	35	0.42
			189	4	37		19	37		45	53	
			191, 193	7	9		8	8		17	11	
			195, 197, 201	5	12		7	17		17	19	
761	35.2	7.7	353, 355	14	49	0.27	28	47	0.89	71	61	0.78
			357, 359	3	22		13	22		31	33	
			361, 363, 365, 371, 373, 375	5	10		7	15		22	20	

P values are shown for comparisons of CEPH rs3849942A and rs3849942G only alleles, all CEPH and case rs3849942A/G heterozygous individuals, and all CEPH and all expansion cases regardless of rs3849942 genotype. Numbers of CEPH and *C9orf72* expansion positive cases used in this analysis were maximised using all available microsatellite data. Microsatellite marker groupings were made with adjacent allele sizes to achieve cell size >5 in each row compatible with chi-Squared test.

Table S2. Details of Microsatellite Haplotyping in CEPH Families which Showed Intergenerational Slippage in Expansion Size

	D9S1814	D9S976	D9S171	D9S1121	D9S169	rs3849942	# G4C2 rpts	D9S263	D9S270	D9S104	D9S147	D9S761
CEPH family 1420												
father	275	138	176	210	272	A	**11**	182	96	190	193	371
	275	136	178	202	264	A	8	180	96	198	195	355
mother	275	128	170	202	272	G	2	180	94	190	191	355
	275	136	170	206	272	G	2	180	94	192	185	355
daughter	275	138	176	210	272	A	**12**	182	96	190	193	371
	275	128	170	202	272	G	2	180	94	190	191	355
CEPH family 1423												
paternal grandfather	277	134	170	202	274	A	**21**	182	96	190	189	351
	277	128	168	182	158	G	4	180	96	192	189	359
paternal grandmother	279	132	176	170	272	A	2	182	96	196	195	359
	277	128	162	206	272	G	7	180	94	198	185	355
father	277	134	170	202	274	A	**22**	182	96	190	189	351
	279	132	176	170	272	G	2	182	96	196	195	359
mother	277	130	170	182	272	A	8	182	96	188	189	355
	275	134	168	182	270	G	6	180	94	188	187	359
son	277	134	170	202	274	A	**20**	182	96	190	189	351
	277	130	170	182	272	A	8	182	96	188	189	355

Details of microsatellite and rs3849942 haplotyping in 2 CEPH families who showed evidence of intergenerational repeat slippage. Haplotypes were confirmed using microsatellite data from entire family to confirm phasing. Transmitted haplotypes are uppermost for each family member and those haplotypes on which slippage occurred are highlighted with the number of repeats in bold type.

Table S3. Detailed Methods Including Probe and Primer Sequences

Technique	Primer Name	Primer Sequence
rs3849942 Genotyping	rs3849942f	ACCTTTCTTCCCACAGGTCTAGC
	rs3849942r	GGGAGATAGATAAGCAAAGGAATAGTGT
	rs3849942A	VIC-TATTTCTTTTATTGCATTTATT-MGB
	rs3849942G	6FAM-TATTTCTTTTGTTGCATTTA-MGB
GGGGCC Genotyping	chr9:27563580F	6FAM-CAAGGAGGGAAACAACCGCAGCC
	chr9:27563465R	GCAGGCACCGCAACCGCAG
Repeat Primed PCR	ALSFTDf_(6FAM)	6FAM-AGT CGC TAG AGG CGA AAG C
	ALSFTDr	TAC GCA TCC CAG TTT GAG ACG GGG GCC GGG GCC GGG GCC GGG G
	ALSFTDanchor	TAC GCA TCC CAG TTT GAG ACG
Southern Blot	C9Probe	DIG-GGGGCCGGGGCCGGGGCCGGGGCCGGGGCC-DIG
Microsatellite Haplotyping	DS91814f	CTTCGATTGCTGGGATTATG
	DS91814r	VIC-GGGCCTGTGAACCTACTGAC
	DS976f	AACCAGTCTTTCTCTGCTAATT
	DS976r	6FAM-GATCCCTGGAACAAACTGAA
	DS171f	AGCTAAGTGAACCTCATCTCTGTCT
	DS171r	NED-ACCCTAGCACTGATGGTATAGTCT
	DS1121f	GTTGTTGCTGGGCTTAGAAA
	DS1121r	VIC-TTGATTGACTGCACTTTCCA
	DS169f	AGAGACAGATCCAGATCCCA
	DS169r	6FAM-TAACAACTCACTGATTATTTAAGGC
	DS263f	TCATTTGGGCAGAGGATCA
	DS263r	PET-TACCTGGGTGGTGACCAGT
	DS270f	AGGTGTAGTCCTTCTGGAATTT
	DS270r	6FAM-GATGTGACTGCTGTTAAAACTAGAG
	DS104f	ACTGGGACTCTAACTAATGT
	DS104r	6FAM-GATCTGGGTATGTCTTTCTG
	DS147Ef	AGTGTTCACCCTAATAAGCC
	DS147Er	NED-CTCCCTGCACCCTTCCATAA
	DS761f	GCCAAGACCACCGCAGTAC
	DS761r	6FAM-GGAAGGGAACCCCCATACG

Details of all primers and probes used in methods described below.

DNA Extraction

Genomic DNA was extracted using the Nucleon BACC2 DNA extraction kit (RPN8502) following the supplied protocol. DNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer, and adjusted to 200-250 ng/µl in TE buffer[2]. Concentrations were re-measured and diluted to 20 ng/µl. Some case samples were extracted from brain tissue as previously described[9].

Microsatellite Analysis

Microsatellite analysis was performed using ten markers spanning approximately 13.1Mb of genomic DNA centred around the *C9orf72* gene. PCR amplicons were generated using fluorescently end labeled primers at 500nM for microsatellite markers D9S1814(VIC), D9S976(FAM), D9S171(NED), D9S1121(VIC), D9S169(FAM), D9S263(HEX), D9S270(FAM), D9S104(FAM), D9S147E(NED) and D9S761(FAM) in MegaMix Royal hot start cocktail (Microzone). Thermal cycling conditions included an initial preheat at 95°C for 5 minutes, followed by 35 cycles of 95°C 30", 58°C 40", 72°C 1'. A loading mix of 1µl amplicon diluted 1:50 in ddH2O, 9.5µl HiDi formamide (ABI) and 0.5µl 500LIZ size standard was prepared and DNA products were electrophoresed on an ABI 3730xl automated sequencer. Data was analysed using ABI GeneMapper software v4.0 (Applied Biosystems (ABI)). To assess the possibility that our haplotype data was incompatible with the single founder hypothesis because of extreme rates of either recombination and/or microsatellite instability, we assessed these attributes in the markers nearest *C9orf72*, D9S263 and D9S169, in the CEPH families. Of 1061 allele transmissions assessed we detected an approximate mutation rate in D9S263 of 1 in 500, and in D9S169 a rate of 1 in 150. Recombination rate between markers was calculated to be approximately 1 in 120, compatible with published rates[12]. Furthermore, our empirical data described above is likely to be an overestimate of microsatellite instability and recombination rates due to the presence of new alleles arising from somatic mutation in CEPH lymphoblastoid cell lines[26].

Rs3849942 Genotyping

The surrogate marker rs3849942, defining the haplotypes at risk of expansion[2], was genotyped by allelic discrimination using the 5' nuclease assay in conjunction with Minor Groove Binding (MGB) probes. The custom designed assay was performed on the SDS7500 Fast Real Time PCR system (ABI) and genotyping calls were made using software v2.0.6.