***GeneReviews* Educational Materials: Comprehensive Genome Sequencing and Multi-Gene Panels**

Supplementary Material: Table 2. Genes and Related Disorders Caused by Nucleotide Repeat Expansions and Contractions

Posted: 14 March 2017

Nucleotide Repeat Expansions and Contractions

A nucleotide repeat is a sequence of n nucleotides repeated a number of times in tandem; nucleotide repeats can occur within or near a gene. The size of nucleotide repeats varies: smaller numbers of repeats are common and not associated with phenotypic abnormalities; abnormally large numbers of repeats may be associated with phenotypic abnormalities and are classified as (in increasing order of size): mutable normal alleles, premutations, reduced-penetrance alleles, and full-penetrance alleles.

Molecular genetic testing used to sequence nucleotide repeats is more difficult than sequencing nonrepetitive regions of the exome because:

- Many of the known nucleotide repeats contain a higher GC content, which is difficult to amplify by PCR; and
- Repetitive regions do not align uniquely; thus, the length of the repeated sequence cannot be determined.

Specific assays are required to analyze each nucleotide repeat of interest:

- DNA containing smaller nucleotide repeats can be amplified by PCR. The amplified segments of DNA are then separated by gel or capillary electrophoresis to determine repeat length.
- Highly expanded nucleotide repeats may not be detected by PCR-based assays due to difficulty in aligning the sequence to a unique genomic position. Additional testing (e.g., Southern blot analysis or triplet repeat primed PCR) may be required to determine the length of highly expanded nucleotide repeats.

Table 2. Genes and Related Disorders Caused by Nucleotide Repeat Expansions and Contractions

Gene	Disorder [1]	% of Pathogenic Variants [2]	Nucleotide Repeat (Amino Acid)	Repeat Location	Normal Repeat Number	Full-Penetrance Pathogenic Repeat Number
AFF2	Fragile X syndrome, FRAXE type (OMIM 309548)	>99%	GCC	5' UTR	6-25	>200
AR	Spinal bulbar muscular atrophy	100%	CAG (Gln)	Exon 1	≤34	≥38
ARX	Early-infantile epileptic	Most common	GCG (Ala)	Exon 2 aa 110-115	10-16	17-27

Gene	Disorder [1]	% of Pathogenic Variants [2]	Nucleotide Repeat (Amino Acid)	Repeat Location	Normal Repeat Number	Full-Penetrance Pathogenic Repeat Number
	encephalopathy (OMIM 308350) Partington syndrome (OMIM 309510)			Exon 2 aa 144-155	12	20
ATN1	Dentatorubral-pallidoluysian atrophy	100%	CAG (Gln)	Exon 5	6-35	≥48
ATXN1	Spinocerebellar ataxia type 1	100%	CAG (Gln)	Exon 8	<35	≥39
ATXN2	Spinocerebellar ataxia type 2	100%	CAG (Gln)	Exon 1	≤31	≥33
ATXN2	L-dopa responsive parkinsonism ALS type 13 (OMIM 183090)	100%	CAG (Gln)	Exon 1	≤31	29-33
ATXN3	Spinocerebellar ataxia type 3	100%	CAG (Gln)	Exon 8	12-44	52-87
ATXN7	Spinocerebellar ataxia type 7	100%	CAG (Gln)	Exon 1	10-19	≥36
ATXN8OS	Spinocerebellar ataxia type 8	100%	CTG	3' UTR	15-50 CTA/CTG	≥71-1300 CTA/CTG
ATXN8	Spinocerebellar ataxia type 8	100%	CAG (Gln)	Exon 1	~80	Unknown
ATXN10	Spinocerebellar ataxia type 10	100%	ATTCT	Intron 9	10-32	≥800
BEAN1	Spinocerebellar ataxia type 31 (OMIM 117210)	100%	TGGAA	Intron 6	0	2.5 to 3.8-kb insertion
C9orf72	Frontotemporal dementia, ALS type 1, hereditary ataxia	5%-30%	GGGGCC (Gly-Ala)	Promotor or intron 1	2-9	700-1600
CACNA1A	Spinocerebellar ataxia type 6	>99%	CAG (Gln)	Exon 7	≤18	20-33
CD40LG	X-linked hyper IgM syndrome	1 family	T	Exon 2	4	5
CNBP	Myotonic dystrophy type 2	100%	CCTG	Intron 1	≤26	>75
COMP	Pseudo-achondroplasia	30%	GAC (Asp)	Exon 13	4	5
CSTB	Unverricht-Lundborg disease	>90%	CCCCGCCCCGCG	5' promoter	2-3	≥30
DIP2B	Mental retardation type 12A (OMIM 136630)	100%	CGG	5' UTR	6-23	>350
DMPK	Myotonic dystrophy type 1	100%	CTG	3' UTR	5-34	>50

Gene	Disorder [1]	% of Pathogenic Variants [2]	Nucleotide Repeat (Amino Acid)	Repeat Location	Normal Repeat Number	Full-Penetrance Pathogenic Repeat Number
EIF4A3	Richieri Costa Pereira syndrome (OMIM 268305)	100%	CACA	5' UTR	5-12	≥15
FMR1	Fragile X syndrome, fragile X-associated tremor/ataxia syndrome	>99%	CGG	5' UTR	5-44	>200
FOXL2	Blepharophimosis, ptosis, epicanthus inversus	31%	GCN (Ala)	Exon 1	14	15-24
FXN	Friedreich ataxia	>90%	GAA	Intron 1	5-33	≥66
HOXA13	Hand-foot-genital syndrome	50%-60%	CCG (Ala)	Exon 1 aa 112 aa 217 aa 346	14 12 12-18	22 18 18-32
HOXD13	Polysyndactyly Syndactyly type V (OMIM 186300)	3 individuals	GCN (Ala)	Exon 1	15	8 ≥22
HTT	Huntington disease	100%	CAG (Gln)	Exon 1	≤26	>40
IL11RA	Craniosynostosis and dental anomalies (OMIM 614188)	1 family	ACCTGGAGC (Thr-Trp-Ser)	Exon 9	2	3
JPH3	Huntington disease-like 2	~100%	CTG/CAG (Ala)	Exon 2A	6-28	≥41
MUC1	Medullary cystic kidney disease	>50%	C	ORF	7	8
MYH7	Laing distal myopathy	5 families	AAG (Lys)	Exon 36	3	2
NOP56	Spinocerebellar ataxia type 36	100%	GGCCTG	Intron 1	3-14	≥650
PABPN1	Oculopharyngeal muscular dystrophy	100%	GCN (Ala)	Exon 1	6	8-13
PAX2	Renal coloboma syndrome	>25%	G	Exon 2	7	6 or 8 or 9
PHOX2B	Congenital central hypoventilation syndrome	97%	GCN (Ala)	Exon 3	≤20	25-29
PPP2R2B	Spinocerebellar ataxia type 12	100%	CAG	5' UTR	8-23	51-78
RUNX2	Cleidocranial dysplasia	1 family	GCN (Ala)	Exon 1	17	27
SOX3	Panhypopituitarism and intellectual disability with growth hormone deficiency (OMIM 300123)	2 families	GCN (Ala)	Exon 1	15	22-26

Gene	Disorder [1]	% of Pathogenic Variants [2]	Nucleotide Repeat (Amino Acid)	Repeat Location	Normal Repeat Number	Full-Penetrance Pathogenic Repeat Number
TBP	Spinocerebellar ataxia type 17	100%	CAG or CAA (Gln)	Exon 3	25-44	47-63
TBX1	Tetralogy of Fallot (OMIM 602054)	1 individual	GCN (Ala)	Exon 9c	15	25
TCF4	Fuchs endothelial corneal dystrophy (OMIM 613267)	~70%	CTG or CAG	Intron 3	<40	>40-50
ZIC2	Holoprosencephaly type 5 (see Holoprosencephaly Overview)	~40%	GCN (Ala)	Exon 3	15	25
ZIC3	VACTERL (OMIM 300265)	1 individual	GCC (Ala)	Exon 1	10	12

The human genome includes >32,000 trinucleotide repeats of ≥6 repeated units. The human exome contains 1030 trinucleotide repeats in exons of 878 genes [Kozlowski et al 2010].
ALS = amyotrophic lateral sclerosis
ORF = open reading frame
UTR = untranslated region
1. For more information see hyperlinked *GeneReview*. An OMIM phenotype entry is provided if a *GeneReview* is not available.
2. Proportion of pathogenic variants in this gene that are caused by a nucleotide repeat expansion or contraction

References

Kozlowski P, de Mezer M, Krzyzosiak WJ. Trinucleotide repeats in human genome and exome. Nucleic Acids Res. 2010;38:4027-39.