Genetics in Medicine (2026) ■, 102520

ELSEVIER

Genetics in Medicine
An Official Journal of the ACMG

www.journals.elsevier.com/genetics-in-medicine

ACMG STATEMENT

Detection of repeat expansion variants using next generation sequencing: A points to consider statement of the American College of Medical Genetics and Genomics (ACMG)

Saurav Guha[1], Indhu-Shree Rajan-Babu[2,3], Akanchha Kesari[4], Shreshtha Garg[5], Nancy C. Rose[6], Bryce A. Seifert[7], Honey V. Reddi[8]; on behalf of the ACMG Laboratory Quality Assurance Committee[9,*]

Disclaimer: This Points to Consider document is designed primarily as an educational resource for clinical laboratory geneticists to help them provide quality clinical laboratory genetic services. Adherence to these Points to Consider is voluntary and does not necessarily assure a successful medical outcome. These Points to Consider should not be considered inclusive of all proper procedures and tests or exclusive of other procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific procedure or test, the clinical laboratory geneticists should apply their own professional judgment to the specific circumstances presented by the individual patient or specimen. Clinical laboratory geneticists are encouraged to document in the patient's record the rationale for the use of a particular procedure or test, whether or not it is in conformance with these Points to Consider. They also are advised to take notice of the date this document was adopted, and to consider other relevant medical and scientific information that becomes available after that date. It also would be prudent to consider whether intellectual property interests may restrict the performance of certain tests and other procedures. Where individual authors are listed, the views expressed may not reflect those of authors' employers or affiliated institutions.

ARTICLE INFO

Article history:
Received 15 January 2026
Accepted 16 January 2026
Available online xxxx

Keywords:
Diagnosis
Next-generation sequencing
Repeat expansions
Screening

Introduction

Short tandem repeats (STRs) are a common type of genetic variation within the human genome and are defined as DNA segments composed of tandemly repeated nucleotide motifs that are usually 2-6 base pairs in length.[1-3] STRs can undergo changes in length, usually an increase in the number of tandem repeats due to various biological processes, making them unstable.[1-3] This instability at specific loci can lead to genetic disorders, collectively known as repeat expansion disorders (REDs).[4] Over 50 REDs, many of which cause neurological disorders, have been associated with expansion of STRs or variable number tandem repeats (VNTRs: ~7-100 bp motifs) (Table 1).[1,2,5-7] The severity and age of onset of a RED varies depending on the size and motif composition of the repeat expansion (RE) and the

The Board of Directors of the American College of Medical Genetics and Genomics approved this statement on 17 November 2025.
*Correspondence: ACMG. *Email address:* documents@acmg.net
Affiliations are at the end of the document.

doi: https://doi.org/10.1016/j.gim.2026.102520
1098-3600/© 2026 American College of Medical Genetics and Genomics. Published by Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

ARTICLE IN PRESS

Table 1 Genetic disorders caused by short tandem repeat expansions and contractions

					Tandem Repeat in GRCh38 Reference Sequence					Pathogenic
Gene/HGNC ID	Disorder	OMIM#	MOI	Pathogenic Repeat Motif (Amino Acid)	GRCh38 Genomic Reference, Genomic Coordinates	Normal Configuration (non-canonical motifs in italics)	Repeat Location	Normal Repeat Number	Intermediate (I), Premutation (P), and Reduced Penetrance (RP) Repeat Number	Repeat Number
AFF2/HGNC:3776	Fragile X syndrome, FRAXE type	309548	XL	CCG	NC_000023.11, chrX:148500606-148500689	CCG *CGG* (CCG)4 *CCT GTG CAG* CCG *CTG* (CCG)15 *CTG* CCG	5′ UTR	≤30	31-60 (I) 61-200 (P)	>200
AR/HGNC:644	Spinal and bulbar muscular atrophy	313200	XL	CAG (Gln)	NC_000023.11, chrX:67545317-67545385	(GCA)23	Exon 1	≤34	36-37 (RP)	≥38
ARX/HGNC:18060	Early-infantile epileptic encephalopathy;	308350	XL	GCG (Ala)	NC_000023.11, chrX:25013650-25013697	(NGC)16	Exon 2 aa 110-115	≤16	-	≥18
	Partington syndrome	309510		GCG (Ala)	NC_000023.11, chrX:25013530-25013565	(NGC)12	Exon 2 aa 144-155	≤12	-	≥20
ATN1/HGNC:3033	DRPLA	125370	AD	CAG (Gln)	NC_000012.12, chr12:6936717-6936773	CAG *CAA* CAG *CAA* (CAG)15	Exon 5	≤35	35-47 (I)	≥48
ATXN1/ HGNC:10548	Spinocerebellar ataxia type 1 Susceptibility to amyotrophic lateral sclerosis	164400	AD	CAG (Gln)	NC_000006.12, chr6:16327636-16327722	(CAG)12 *CAT* CAG *CAT* (CAG)14	Exon 8	≤35	36-38u (I) 44i (RP)	39-44i or 46-70i/u
ATXN2/ HGNC:10555	Spinocerebellar ataxia type 2 Susceptibility to amyotrophic lateral sclerosis 13 and Parkinsonian phenotype	183090	AD	CAG (Gln)	NC_000012.12, chr12:111598951-111599019	(CAG)13 *CAA* (CAG)9	Exon 1	≤31	33-34 (RP)	≥37
ATXN3/ HGNC:7106	Spinocerebellar ataxia type 3	109150	AD	CAG (Gln)	NC_000014.9, chr14:92071011-92071052	(CAG)2 *CAA AAG* CAG *CAA* (CAG)8	Exon 10	≤44	45-60 (I)	~60-87
ATXN7/ HGNC:10560	Spinocerebellar ataxia type 7	164500	AD	CAG (Gln)	NC_000003.12, chr3:63912686-63912715	(CAG)10	Exon 1	≤27	34-36 (RP)	37-460
ATXN8/ HGNC:32925	Spinocerebellar ataxia type 8	608768	AD	CAG (Gln)	NC_000013.11, chr13:70139384-70139428	(CAG)15	Exon 1	-	-	-
ATXN8OS/ HGNC:10561				CTG		(CTG)15	3′ UTR	≤50 CTA/CTG	[a]	[a]
ATXN10/ HGNC:10549	Spinocerebellar ataxia type 10	603516	AD	ATTCT	NC_000022.11, chr22:45795355-45795424	(ATTCT)14	Intron 9	≤32	~33-850 (RP)	≥800
BEAN1/ HGNC:24160	Spinocerebellar ataxia type 31	117210	AD	TGGAA	NC_000016.10, chr16:66490399-66490466	(TAAAA)11	Intron 6	TAAAA? TGGAA 0	-	≥500
C9orf72/ HGNC:28337	C9orf72-related amyotrophic lateral sclerosis and frontotemporal dementia	*105550*	AD	GGGGCC	NC_000009.12, chr9:27573529-27573546	(GGGGCC)3	Promotor or intron 1	≤24	25-60 (uncertain significance)	>60
CACNA1A/ HGNC:1388	Spinocerebellar ataxia type 6	183086	AD	CAG (Gln)	NC_000019.10, chr19:13207859-13207897	(CAG)13	Exon 47	≤18	19 (uncertain significance)	20-33

(continued)

ARTICLE IN PRESS

Table 1 Continued

					Tandem Repeat in GRCh38 Reference Sequence					Pathogenic
Gene/HGNC ID	Disorder	OMIM#	MOI	Pathogenic Repeat Motif (Amino Acid)	GRCh38 Genomic Reference, Genomic Coordinates	Normal Configuration (non-canonical motifs in italics)	Repeat Location	Normal Repeat Number	Intermediate (I), Premutation (P), and Reduced Penetrance (RP) Repeat Number	Repeat Number
CNBP/ HGNC:13164	Myotonic dystrophy type 2	602668	AD	CCTG	NC_000003.12, chr3:129172577-129172656	(CCTG)8 *GCTG* CCTG *TCTG* CCTG *TCTG* (CCTG)7	Intron 1	≤26	~30-54 (P) ~55-74 (uncertain significance)	≥75
COMP/HGNC:2227	Multiple epiphyseal dysplasia	132400	AD	GAC (Asp)	NC_000019.10, chr19:18786035-18786049	(GAC)5	Exon 13	5	-	≥6
	Pseudoachondroplasia	177170							-	2-4 or ≥7
CSTB/HGNC:2482	Progressive myoclonic epilepsy type 1	254800	AR	CCCCGCCCCGCG	NC_000021.9, chr21:43776444-43776479	(CCCCGCCCCGCG)3	Promoter	≤3	-	≥30
DAB1/HGNC:2661	Spinocerebellar ataxia type 37	615945	AD	ATTTC	NC_000001.11, chr1:57367044-57367118	(ATTTT)15	5′ UTR intron	ATTTT? ATTTC 0	-	31-75
DIP2B/ HGNC:29284	Intellectual developmental disorder, FRA12A type	136630	AD	CGG	NC_000012.12, chr12:50505004-50505024	(CGG)7	Promoter	≤23	-	>350
DMPK/HGNC:2933	Myotonic dystrophy type 1	160900	AD	CTG	NC_000019.10, chr19:45770205-45770264	(CTG)20	3′ UTR	≤34	35-49 (P)	>50
EIF4A3/ HGNC:18683	Pierre Robin sequence with cleft mandible and limb anomalies	268305	AR	Complex[b]	NC_000017.11, chr17:80147060-80147139	(TCGGCAGCGGCA CAGCGAGG)4	5′ UTR	≤12	-	≥14
FGF14/HGNC:3671	Spinocerebellar ataxia type 27B	620174	AD	GAA	NC_000013.11, chr13:102161577-102161726	(GAA)50	Intron 1	≤249	250-335 (RP)	>300
FMR1/HGNC:3775	FMR1-related disorders	300624, 300623, 311360	XL	CGG	NC_000023.11, chrX:147912051-147912110	(CGG)10 *AGG* (CGG)9	5′ UTR	≤44	45-54 (I) 55-200 (P)	>200
FOXL2/HGNC:1092	Blepharophimosis, ptosis, and epicanthus inversus	110100	AD	GCN (Ala)	NC_000003.12, chr3:138946021-138946062	(GCN)14	Exon 1	≤14	Undefined	19-24
FXN/HGNC:3951	Friedreich ataxia	229300	AR	GAA	NC_000009.12, chr9:69037287-69037304	(GAA)6	Intron 1	≤33	34-65 (P)	≥66
GIPC1/HGNC:1226	Oculopharyngodistal myopathy 2	618940	AD	GGC	NC_000019.10, chr19:14496042-14496083	CGG *AGG CAG* (CGG)11	5′ UTR	≤32	-	73-164
HOXA13/ HGNC:5102	Hand-foot-genital syndrome	140000	AD	GCN (Ala)	NC_000007.14, chr7:27199925-27199966	(GCN)14	Exon 1 aa 38	≤14	-	22
					NC_000007.14, chr7:27199826-27199861	(GCN)12	Exon 1 aa 73	≤12	-	18
					NC_000007.14, chr7:27199679-27199732	(GCN)18	Exon 1 aa 116	≤18	-	24-32
HTT/HGNC:4851	Huntington disease	143100	AD	CAG (Gln)	NC_000004.12, chr4:3074877-3074966	(CAG)19 *CAA* CAG CCG *CCA* (CCG)7	Exon 1	≤26	27-35 (I) 36-39 (RP)	≥40
JPH3/ HGNC:14203	Huntington disease-like 2	606438	AD	CTG (Ala)	NC_000016.10, chr16:87604287-87604329	(CTG)14	Exon 2A	≤28	29-39 (uncertain significance)	≥40

(continued)

Table 1 Continued

Gene/HGNC ID	Disorder	OMIM#	MOI	Pathogenic Repeat Motif (Amino Acid)	Tandem Repeat in GRCh38 Reference Sequence: GRCh38 Genomic Reference, Genomic Coordinates	Tandem Repeat in GRCh38 Reference Sequence: Normal Configuration (non-canonical motifs in italics)	Repeat Location	Normal Repeat Number	Intermediate (I), Premutation (P), and Reduced Penetrance (RP) Repeat Number	Pathogenic Repeat Number
LRP12/ HGNC:31708	Oculopharyngodistal myopathy Susceptibility to amyotrophic lateral sclerosis	164310	AD	CGG/CGT	NC_000008.11, chr8:104588960-104588998	(CGG)9 *CGT* CGG (*CGT*)2	5′ UTR	≤45	-	≥93
MARCHF6/ HGNC:30550	Familial adult myoclonic epilepsy 3	613608	AD	TTTCA	NC_000005.10, chr5:10356347-10356411	(ATTTT)13	Intron 1	ATTTT ≤20 TTTCA 0	-	791-1035
MUC1/HGNC:7508	Autosomal dominant tubulointerstitial kidney disease, MUC1-related	174000	AD	C[c]	NC_000001.11, chr1:155188487-155192239	-	Exon 2	≤7	-	8
NOP56/ HGNC:15911	Spinocerebellar ataxia type 36	614153	AD	GGCCTG	NC_000020.11, chr20:2652734-2652757	(GGCCTG)4	Intron 1	≤14	-	≥650
NOTCH2NLC/ HGNC:53924	Neuronal intranuclear inclusion disease	603472	AD	GGC	NC_000001.11, chr1:149390803-149390841	(GGC)9 *GGA GGA* (GGC)2	5′ UTR	≤39	-	≥66
	Oculopharyngodistal myopathy 3	619473								>100
	Hereditary essential tremor 6	618866						≤41		≥47
PABPN1/ HGNC:8565	Oculopharyngeal muscular dystrophy	164300	AD	GCN (Ala)	NC_000014.9, chr14:23321473-23321490	(GCG)6	Exon 1	≤10	-	11-18
PHOX2B/ HGNC:9143	Congenital central hypoventilation syndrome	209880	AD	GCN (Ala)	NC_000004.12, chr4:41745972-41746031	(GCN)20	Exon 3	≤20	20-24 (RP or incomplete penetrance)	≥24
PPP2R2B/ HGNC:9305	Spinocerebellar ataxia type 12	604326	AD	CAG	NC_000005.10, chr5:146878728-146878757	(CAG)10	Promoter	≤31	-	≥43
PRNP/HGNC:9449	Creutzfeldt-Jakob disease	123400	AD	CCTCATGGTGGTG GCTGGGGGCAG	NC_000020.11, chr20:4699371-4699493	*CCTCAGGGCGGTGGTG GCTGGGGGCAG* (CCTCATGGTGGT GGCTGGGGGCAG)2 *CCCCATGGTGGTG GCTGGGGGCAG* CC TCATGGTGGTG GCTGGGGTCAA	Exon 2	≤4	-	5-16
RAPGEF2/ HGNC:16854	Familial adult myoclonic epilepsy type 7	618075	AD	TTTCA	NC_000004.12, chr4:159342527-159342616	(TTTTA)5 *TATTA* (TTTTA)12	Intron 14	TTTTA? TTTCA 0	-	≥60
RFC1/HGNC:9969	RFC1 CANVAS/ spectrum disorder	614575	AR	AAGGG	NC_000004.12, chr4:39348425-39348479	(AAAAG)11	Intron 2	AAAAG or AAAGG 11-1000 AAGGG 0	AAAGG (variable penetrance)	400 to >2000
RILPL1/ HGNC:26814	Oculopharyngodistal myopathy 4	619790	AD	GGC	NC_000012.12, chr12:123533721-123533750	(GGC)10	5′ UTR	≤16	-	≥135
RUNX2/ HGNC:10472	Cleidocranial dysplasia spectrum disorder	119600	AD	GCN (Ala)	NC_000006.12, chr6:45422751-45422801	(GCN)17	Exon 1	17	-	20-27
SAMD12/ HGNC:31750	Familial adult myoclonic epilepsy type 1	601068	AD	TTTCA/TTTGA	NC_000008.11, chr8:118366816-118366918	(TTTTA)7 TTA (TTTTA)13	Intron 4	TTTTA? TTTCA or TTTGA 0	-	7-14 or ≥100

(continued)

ARTICLE IN PRESS

ARTICLE IN PRESS

Table 1 Continued

Gene/HGNC ID	Disorder	OMIM#	MOI	Pathogenic Repeat Motif (Amino Acid)	Tandem Repeat in GRCh38 Reference Sequence: GRCh38 Genomic Reference, Genomic Coordinates	Tandem Repeat in GRCh38 Reference Sequence: Normal Configuration (non-canonical motifs in italics)	Repeat Location	Normal Repeat Number	Intermediate (I), Premutation (P), and Reduced Penetrance (RP) Repeat Number	Pathogenic Repeat Number
TBP/HGNC:11588	Spinocerebellar ataxia type 17	607136	AD	CAG or CAA (Gln)	NC_000006.12, chr6:170561908-170562017	(CAG)3 *CAA CAA CAA* (CAG)8 *CAA* CAG *CAA* (CAG)19	Exon 3	≤40	41-48 (RP)	≥49
TCF4/HGNC:11634	Fuchs endothelial corneal dystrophy 3	613267	AD	CTG or CAG	NC_000018.10, chr18:55586156-55586227	(CTG)24	Intron 3	≤39	-	≥50
THAP11/ HGNC:23194	Spinocerebellar ataxia subtype 51	620947	AD	CAG (Gln)	NC_000016.10, chr16:67842864-67842950	(CAG)3 *CAA* (CAG)5 *CAA* (CAG)2 *CAA* (CAG)5 *CAA* (CAG)10	Exon 1	≤38	-	45-100
VWA1/ HGNC:30910	Hereditary motor neuropathy	619216	AR	GGCGCGGAGC	NC_000001.11, chr1:1435799-1435818	(GGCGCGGAGC)2	Exon 1	2	-	1 or 3
XYLT1/ HGNC:15516	Baratela-Scott syndrome (Desbuquois dysplasia type 2)	615777	AR	GGC	NC_000016.10, chr16:17470908-17470922	(GGC)5	Promoter	≤20	-	~>72
ZFHX3/HGNC:777	Spinocerebellar ataxia type 4	600223	AD	GGC (Gly)	NC_000016.10, chr16:72787695-72787757	(GCC)8 *ACC* (GCC)5 *ACT* GCC *ACC* (GCC)4	Exon 10	≤26	-	46-64
ZIC2/HGNC:12873	Holoprosencephaly type 5	609637	AD	GCN (Ala)	NC_000013.11, chr13:99985449-99985493	(GCN)15	Exon 3	≤15	-	25

The list is not exhaustive and limited to the REDs with established genotype/phenotype correlations. The table has been compiled based on Rajan-Babu et al,[3] 2024.

aa, amino acid; AD, autosomal dominant; *AR*, autosomal recessive; *i*, interrupted by CAT repeat motifs; *MOI*, mode of inheritance; *u*, uninterrupted or absence of CAT repeat motifs; *UTR*, untranslated region; *XL*, X-linked.

[a]Penetrance is <100%. Increased penetrance is reported for alleles of 54-250 CTA/CTG repeats.

[b]This repeat comprises repeating units of 18 or 20 nucleotides that vary at a CA sequence.

- Normal repeat: CACA-20-nt(2-9)CA-18-nt(1)CACA-20-nt(1)CA-18-nt(1), note, a normal allele has 5-12 total repeats.
- Abnormal allele: CACA-20-nt(1) CGCA-20-nt(12-13)CA-18-nt(1)CACA-20-nt(1)CA-18-nt(1), note, a normal allele has 15-16 total repeats.

[c]Duplication of 1 cytosine in a heptanucleotide cytosine tract within 1 copy of a 20-125 copy-number VNTR (variable number tandem repeat). The specific VNTR involved varies by family but is consistent within a family.

function of that gene (Table 1). Two main types of REDs are recognized. The first type affects coding regions of genes, often resulting in abnormally long stretches of specific amino acids, such as polyglutamine (polyQ), polyalanine (polyA), or polyaspartic acid (polyD), within proteins. The second type involves REs in noncoding regions of genes, including the 5′ or 3′ untranslated regions or introns, where they can disrupt gene expression or RNA processing.[3] Pinpointing the exact pathophysiologic mechanism is often challenging because several may be occurring simultaneously.[1-3,5] Broadly, REDs cause disease through loss-of-function and/or gain-of-function mechanisms, with both potentially contributing in some cases.[1,3,5]

The size of the RE generally correlates with disease severity: the larger the RE, the earlier and more severe the symptoms. However, the threshold at which these REs become pathogenic is not uniform and varies by disorder.[1,2,5] Alleles in REDs are typically classified as normal, intermediate (premutation), or pathogenic (full mutation) range as defined by gene-specific guidelines. For example, in *FMR1* (HGNC:3775): normal ≤ 44 repeats, intermediate 45-54 repeats, premutation 55-200 repeats, and pathogenic ≥ 200. An intermediate or premutation allele, defined by repeat lengths between the established normal and full mutation ranges, may confer an increased risk for conditions beyond the classic RED and may also exhibit instability during parent-to-child transmission.[1-3] Accurately detecting the number and length of the REs is essential for diagnosis and genetic counseling, given that genetic anticipation (increase in RE size from one generation to the next) is common in REDs. Historically, the detection of REs involved techniques such as long-range PCR, repeat-primed PCR (RP-PCR), and/or Southern blot assays (Table 2).[2,8-10] Despite their higher sensitivity, these techniques are resource-intensive and have limitations, such as the need for gene-specific primers/probes and the inability to analyze multiple loci simultaneously.[6] This presents a challenge for comprehensive diagnostic testing, as many new REDs have been identified in recent years[2] (Table 1). Furthermore, many REDs involve complex repeat structures and overlapping phenotypes, making it exceedingly difficult for laboratories to develop, validate, and standardize specific RP-PCR and/or Southern blot assays for each locus. Variable expressivity is observed in certain REDs; for example, *ATXN2* (HGNC:10555) expansions are associated with amyotrophic lateral sclerosis (ALS, OMIM 183090) risk or spinocerebellar ataxia type 2 (OMIM 183090), depending on the number of CAG (poly Q) repeats and the presence or absence of a CAA interruption, supporting the notion that overlapping phenotypes may be influenced by the number and structure of expanded repeats. In addition, variations in age of onset and disease penetrance are influenced by allele size, repeat interruptions, mosaicism, methylation, and other genetic or epigenetic modifiers, making it challenging to provide accurate counseling regarding phenotypes and long-term outcomes.[1-3,5,11]

Exome and genome sequencing using short-read based next-generation sequencing (NGS) is now a standard test for patients with a suspected genetic disorder (Table 2).[12,13] Although primarily designed to detect single-nucleotide variants, insertions and deletions, and copy-number variants, recent advancements in bioinformatics tools enable NGS to analyze multiple RE loci simultaneously and identify REs in genes of interest.[14-18] Although some tools can identify REs in both exome and genome data, others are tailored specifically for genomes. Several recent studies have demonstrated the utility of exome/genome sequencing in diagnosing REDs.[19-21] However, current short-read NGS methods for RE detection continue to face challenges due to both technological limitations and the complex nature of these loci but are widely implemented across clinical laboratories worldwide.[15,16] Methodologies for detecting REs continue to advance with the development of long-read sequencing technologies, which enable more accurate and comprehensive analysis of REs and hold significant potential for broader future adoption in the molecular diagnosis of REDs.[3]

The American College of Medical Genetics and Genomics (ACMG) has published several Standards and Guidelines for Clinical Genetics Laboratories regarding the detection of RE in specific REDs using RP-PCR and/or Southern blot analysis.[8-10] Given the increased use of NGS for the detection of REs, this Points to Consider document addresses the complexities specifically associated with short-read NGS-based RE testing to facilitate a better understanding for the end user. It outlines the emerging techniques for detecting REs in clinical settings, provides considerations for clinical laboratories offering or introducing RE testing through NGS, and offers considerations for health care providers regarding comprehensive RE analysis from exome and genome data to facilitate the detection of clinically relevant disorders.

Materials and Methods

Review of the current literature and available guidelines, including PubMed and relevant ACMG, Association for Molecular Pathology, College of American Pathologists, and American Society of Clinical Oncology guidelines was performed. The workgroup members also used their expert opinion and empirical data to inform their recommendations. The ACMG Laboratory Quality Assurance Committee reviewed the document providing further input on the content, and a final draft was presented to the ACMG Board of Directors for review and approval to post on the ACMG website for member comment. Upon posting to the ACMG website, an email and link were sent to all ACMG members inviting participation in the 30-day open comment process. All members' comments and additional evidence received were assessed by the authors, and these recommendations were incorporated into the document as

Table 2 Short tandem repeat analysis approaches

		Repeat Length Analysis			Sequence Composition Analysis				
Approaches	Parallel Multi-Loci Analysis	Complex Loci	Determination of Repeat Length	Genotyping Limit	Identification and Quantification of Noncanonical Motifs	Determination of the Sequence of Noncanonical Motifs	GC-Bias	Repeat Length Mosaicism Analysis	Methylation Analysis
Repeat-Primed Polymerase Chain Reaction	N	N	Y	~200 triplet repeats	Y (Limited)[a]	N[b]	Y	Y	Y[c]
Sanger Sequencing	N	Y	Y	~250-300 triplet repeats	Y (Limited)	Y	Y	N	Y[c]
Southern Blot Analysis	N	Y	Y	N	N	N	N	Y	Y
Next-Generation (Short-Read) Sequencing									
Targeted Gene Panels	Y	Y (Limited)	Y	≥RL or FL	Y (RL)	Y (RL)	Y	N	N
Exome	Y	Y (Limited)	Y	≥RL or FL	Y (RL)	Y (RL)	Y	N	N
PCR-Free Whole Genome	Y	Y (Limited)	Y	≥RL or FL	Y (RL)	Y (RL)	N (some)	N	N
Next-Generation (Long-Read) Sequencing									
Targeted (PCR-Based)	Y	Y	Y	N	Y	Y	Y	Y	N
Targeted (Crispr-Cas9)	Y	Y	Y	N	Y	Y	N	Y	Y
Targeted (Adaptive Sampling)	Y	Y	Y	N	Y	Y	N	Y	Y
Whole Genome	Y	Y	Y	N	Y	Y	N	Y	Y
Optical Genome Mapping	Y	Y	Y	<~500 base pairs	N	N	N	Y	N

FL, fragment length (typically 350-650 base pairs); *GC*, guanine-cytosine, *N*, no; *RL*, read length (typically 150-250 base pairs); *Y*, yes.

[a]For expansions >~100 triplet repeat units, depending on the assay design (unidirectional: 5′ or 3′) or bidirectional (5′ and 3′), repeat-primed polymerase chain reaction can only identify noncanonical motifs at one or both flanks of the expanded repeat tract.

[b]Determination of the sequence of the non-canonical motifs is not possible with repeat-primed polymerase chain reaction with the exception of nonreference expansions.

[c]Possible with special methylation-specific protocols that involve bisulfite treatment of genomic DNA or treatment of DNA with methylation-sensitive restriction enzyme before amplification.

ARTICLE IN PRESS

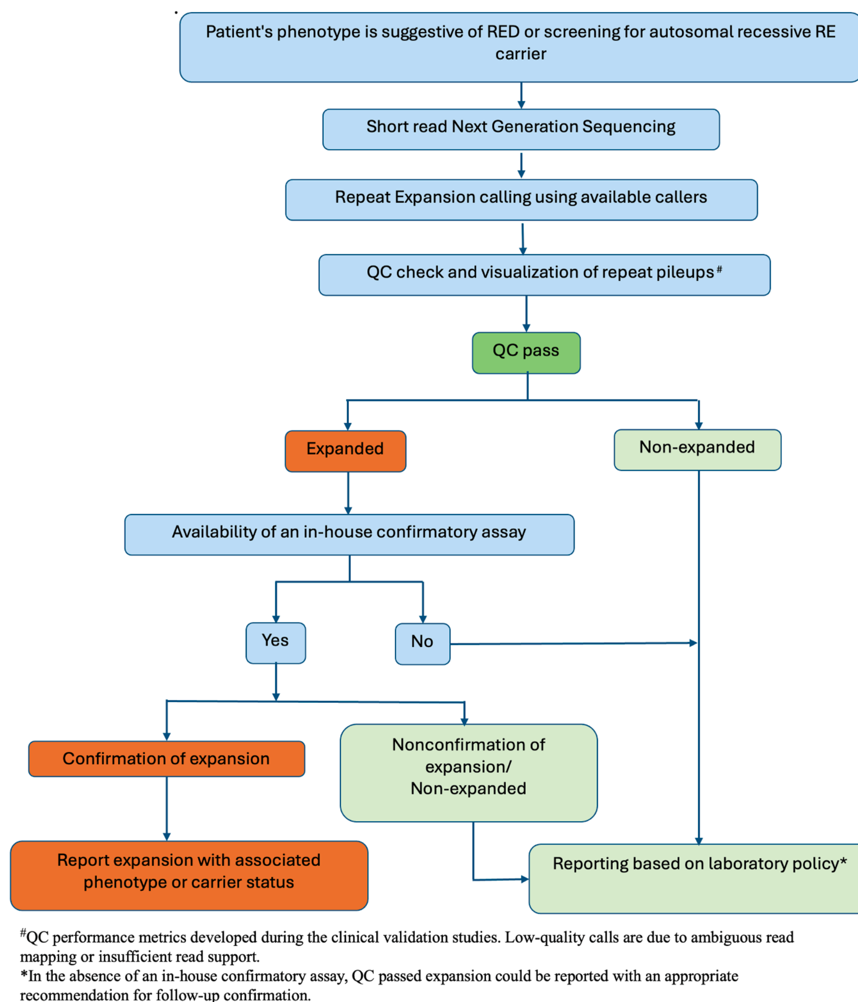

Figure 1 Laboratory flow chart for confirmation and reporting of RE loci using short-read NGS.

deemed appropriate. Member comments and author responses were reviewed by representatives of the ACMG Laboratory Quality Assurance Committee and the ACMG Board of Directors. The final document was approved for publication by the ACMG Board of Directors.

General considerations

- Repeat expansion size and disease associations: Accurate sizing of the RE is critical for diagnosis because the disease phenotype often correlates with repeat length and risk of expansion (anticipation). Certain REs are associated with specific genetic disorders, such as Huntington's disease (OMIM 143100) (CAG repeats in *HTT*, HGNC:4851) or fragile X syndrome (OMIM 300624) (CGG repeats in *FMR1*). The clinical context is essential to focus on relevant loci. Many RE loci have defined size thresholds for pathogenicity, and it is important to distinguish pathogenic from non-pathogenic alleles (Table 1).
- Somatic mosaicism: REs can vary in length across different tissues because of somatic mosaicism, which complicates detection. Sampling different tissues (eg, blood or saliva) may yield different outcomes, especially in diseases like myotonic dystrophy type 1 (DM1, OMIM 160900).[22]
- Interrupted repeat alleles: REs often contain different motifs embedded within their sequence, some of which have been known to play a crucial role in modulating the stability and pathology associated with REDs. Interrupted repeat alleles have been associated with the stabilization of repeat loci and modified clinical presentation.[23,24] For instance, repeat interruptions composed of CCG and GGC motifs within DMPK (HGNC:2933) CTG REs modify mutational dynamics and likely contribute to milder or unusual disease progression or unknown clinical implications and/or interpretation[3,25] (Table 1).
- Incomplete Penetrance and Variable Expressivity: REs may not always lead to clinical findings in every individual, and the severity of the phenotype can vary widely. This complicates genetic counseling and risk assessment (Table 1).
- Parental Testing and Anticipation: Because REs can increase in size in subsequent generations, testing

parents or multiple family members may help assess risk for future offspring within a family.

- Technical limitations: Standard PCR methods may fail to amplify large REs, necessitating alternative techniques, such as RP-PCR or Southern blotting for very large REs. Short-read based NGS is a powerful method for detecting REs; however, short-read based NGS results for REs, especially large ones, should ideally be confirmed with alternative methods, such as Southern blot, RP-PCR, or long-read sequencing based NGS, especially when used for clinical diagnostics (Table 2).
- Clinical utility of testing: Consideration of the clinical utility of RE testing, particularly in the context of conditions for which early diagnosis can lead to improved management or treatment options in disorders such as spinocerebellar ataxias or fragile X syndrome.

Considerations for testing laboratories

A variety of approaches, including molecular assays and genome sequencing/mapping technologies, are available to detect and characterize REDs, as summarized in Table 2. These approaches offer different utilities and present unique advantages and challenges in genotyping, such as determining the accurate length or size of the RE, mapping the sequence composition of the RE, and/or detecting any associated DNA methylation changes. Clinical NGS can enable screening for REDs, with improvements in sequencing and repeat analytical tools now making it feasible to analyze multiple loci in parallel.

Technical considerations

Choice of sequencing technology

- The ability to reliably detect and genotype REDs by NGS depends on the methodology used for library preparation, the sequencing read and DNA fragment length, sequencing depth, and the ability of the STR callers to leverage the different types of reads that originate from repeat regions.
- Exome sequencing enables the analysis of STR loci located within enriched genes or coding sequences, whereas genome sequencing enables the analysis of STRs in both coding and noncoding genomic regions.
- Inherent biases in the chosen technology can affect the detection of some REs. For example, PCR amplification of the *FMR1* locus is adversely affected depending on the number of repeats.[20,22] In general, PCR-free genome sequencing, which offers a more uniform coverage of the genome, yields the best data for detecting REDs. Therefore, PCR-free genome sequencing is recommended for optimal detection of REs in noncoding regions and REs in regions prone to coverage bias.

Choice of analysis pipelines

- The performance of STR analysis tools for RE detection varies based on the types of reads used—tools that rely solely on spanning reads are limited to detecting REs shorter than ~150 to 250 bp, whereas those integrating spanning, flanking, and in-repeat reads perform better for REs >350-550 bp.[17,26] Although spanning reads provide more accurate genotypes, confidence intervals are typically reported for REs near or beyond read length. Additionally, software performance can be influenced by the aligner used; thus, pipelines should be validated if alternate aligners are used. An ensemble approach using multiple tools enhances confidence in RE calls and reduces false positives, whereas adjusting the thresholds for confirmation (eg, lowering cutoffs for undersized loci, such as *FMR1*) may help reduce false negatives.[16]
- Most STR genotyping tools rely on predefined repeat catalogs that limit analysis to specified loci, with the number and content of loci varying by tool. Users may add missing loci, although this is complex for repeats with intricate motifs and requires precise structural knowledge of the repeat structure. Repeat catalogs are critical because even minor coordinate discrepancies can affect repeat length estimates. Furthermore, to integrate in-repeat reads effectively, catalogs should include off-target loci (other STRs with the same motif where reads may map incorrectly), which can improve accuracy but may also increase false positives.[16,17]
- Visual inspection of candidate REs using specialized visualization software is valuable for ruling out false positives, identifying repeat motif changes across the full length of REs or at the distal ends of large expansions, and mapping interruptions at both the 5′ and 3′ ends—capabilities that surpass those of RP-PCR, which only targets one end and displays all interruptions as gaps. Such software also enables characterization of interrupting motifs and allows haplotype-resolved analysis to distinguish interruption patterns between alleles, which can be challenging due to overlapping peaks from both in RP-PCR profiles.[27,28]

Validation of analysis pipelines

The rapid pace of disease-associated STR loci discovery necessitates the maintenance of catalogs for RE analysis by software developers and/or users and the validation of bioinformatics pipelines using control samples with REs of

interest. Often, obtaining access to published data (FASTQ/BAM files) for pipeline optimization is not possible because of the ethical constraints that limit the open sharing of identifiable genomic sequences of individuals. Simulated data sets are an alternative to real experimental data and have been used in several studies to develop and validate computational methods. However, care should be exercised in concluding pipeline/tool performances exclusively based on simulations because simulated data do not faithfully reflect the inherent, systematic biases of sequencing, which could result in erroneous inflation of the performance of tested tools. For instance, the superior performance of RE genotyping software on simulated *FMR1* REs is because of the improved and uniform coverage of the GC-rich sequences in simulated compared with real PCR+ and PCR-free GS.[16]

Control data sets

Resources and databases of benign, known pathogenic repeats, their allele frequencies and literature are available for public use, eg, STRipy, population data from the 1000 Genomes Project (https://stripy.org/),[29] gnomAD STR database (https://gnomad.broadinstitute.org/short-tandem-repeats?dataset=gnomad_r4), Tandem Repeat Aggregation Atlas (TR-Atlas) (https://wlcb.oit.uci.edu/TRatlas/),[30] and STRchive (https://strchive.org/).[7] Positive controls from EGAS00001002462, EGAS00001002598, and EGAS00001002593 are available upon request to tweak and optimize pipelines; however, these resources lack samples with specific disease-associated REs.[31] Control cell lines are available for known REDs from the Centers for Disease Control and Prevention Genetic Testing Reference Materials Coordination Program (https://www.cdc.gov/lab-quality/php/get-rm/index.html) and the Coriell Cell Repositories (https://www.coriell.org/). However, no control samples for any of the pentanucleotide REs are available, highlighting the need for more control cell lines to be established to ensure that the clinical labs can test, optimize, and validate their workflows effectively.

Confirmation of REs

Confirmation of RE loci detected by short-read based NGS by an alternative method is essential due to the technical challenges in sizing REs with NGS alone. Techniques such as RP-PCR and Southern blotting are commonly used to confirm REs identified by NGS, providing higher sensitivity and precision in determining repeat sizes and interruptions.[32] Alternative confirmations are essential for clinical diagnostics because they mitigate the risk of false positives and false negatives and confirm allele length and composition, essential for accurate interpretation of REs.[33] The working group recognizes that the ability of a confirmatory technique to reduce the risk of false negatives depends on its consistent use when normal or nonpositive results are obtained. It is therefore recommended that laboratories establish defined cutoffs to guide when confirmatory testing should be performed.

Recommendations for short-read NGS-based RE detection

- It is recommended that the laboratory develop performance metrics focused on validating the analytical performance of the RE, ensuring its ability to accurately and reliably detect the expanded and non-expanded repeats in all categories (normal, premutation, and full mutation). Laboratories must carefully consider these metrics during assay design, validation, and routine operation to meet the required standards and ensure the quality of their NGS testing services is robust and reliable.
- Whenever possible, laboratories may want to perform confirmatory testing of expanded alleles using an alternative methodology as a standard component of the testing workflow for RE loci (Figure 1), using set RE cutoffs for when confirmatory testing needs to be performed based on laboratory validation metrics.
- In cases in which a confirmatory assay is unavailable, laboratories may report a RE with an appropriate recommendation for follow-up confirmation if indicated.
- As is true with any other molecular method, in the absence of appropriate proficiency testing products available from the College of American Pathologists (CAP) or other vendors for the variant type in question, it is recommended that laboratories use alternative assessment approaches instead in order to assess the accuracy of their method for the detection of any changes associated with that variant type.
- Comprehensive prospective studies comparing the cost and efficacy of the NGS-based methods for screening for REs should also be performed before these approaches are widely adopted.

Additional emerging technologies

Long-read sequencing

- Long-read sequencing (~15-20 kb reads) can genotype REs and detect mosaicism, interruptions, and methylation changes in a single assay; targeted approaches (eg, CRISPR-Cas9 enrichment and adaptive sampling) enable parallel analysis of all disease-associated STR loci (Table 2).[34]
- With improving accuracy and declining costs, long-read sequencing may serve as a valuable follow-up to short-read RE screening, complementing existing clinically validated tests and enhancing genotype-phenotype correlation analyses for better understanding and treatment of RE disorders, as well as more precise genetic counseling.[34]

ARTICLE IN PRESS

- Limitations include the need for high-molecular-weight DNA, elevated sequencing and data analysis costs, and higher error rates compared with short-read sequencing—especially for reliably determining RE motif composition—necessitating further technological refinements and benchmarking. Although the accuracy of long-read sequencing has improved significantly, challenges remain in reliably resolving repeat motif composition under suboptimal coverage conditions. These challenges are technology-specific and are being addressed through ongoing platform improvements.

Optical genome mapping (OGM)

- OGM offers high-resolution detection of large genomic rearrangements and is increasingly used in rare disease diagnostics, especially for large REs, such as those in Huntington's disease or fragile X syndrome, because of its direct visualization of repeats without amplification bias.[35]
- Despite its advantages, OGM has limitations: it lacks sensitivity for small REs, cannot detect single-nucleotide variants or differentiate pathogenic from nonpathogenic motifs (eg, in *RFC1*, HGNC:9969), and cannot identify repeat interruptions.
- The method requires high-quality, high-molecular-weight DNA, has lower throughput than NGS, and lacks clinical standardization, with complex data interpretation needing specialized tools and expertise; yet, it remains a promising approach for analyzing large REs and structural variants.[35,36]

Reporting considerations

Clinical reports for REs should follow the standard regulatory requirements for reporting clinically significant results and should at a minimum include the following sections. For asymptomatic testing, laboratory reports for asymptomatic individuals should not include predictions of symptoms or age of onset based on expansion size (see "Test ordering and report evaluation" for details).

- Reporting of NGS-based RE alleles can aid incorporation of RE alleles into the electronic health record. If incorporating alleles into their electronic health record, laboratories should ensure alleles are notated using repeated sequences variant nomenclature guidelines (https://hgvs-nomenclature.org/stable/recommendations/DNA/repeated/) as set forth by the Human Genome Variation Society.[37]
- Report should include zygosity and the number of repeats in accordance with Human Genome Variation Society repeated sequences variant nomenclature guidelines.
- In accordance with state-accreditation and/or CAP-accreditation policies, laboratories should clearly list the range of all interpretable repeat types (eg, normal, intermediate/premutation, reduced penetrance, full mutation, etc) on reports (for reference, the applicable CAP checklist item is MOL.36147). Alleles should be reported as normal, intermediate (premutation), or pathogenic (full mutation) range as defined by gene-specific guidelines. For example, in *FMR1*: normal $\leq$ 44 repeats, intermediate 45-54 repeats, premutation 55-200 repeats, pathogenic $\geq$ 200 repeats.
- For reporting of in-house laboratory-developed tests, limit of detection and reporting threshold must be determined by individual laboratories. Additionally, laboratories need to determine when confirmation by alternative methods is necessary (eg, RP-PCR, Southern blot, long-read sequencing, etc).
- Reporting of specific repeat numbers is recommended only when the call is evaluated by an alternative method or according to the established laboratory policies for confirmatory testing. In the absence of confirmatory testing, a specific repeat number may be reported when the call is supported by spanning reads because these provide the highest confidence. Otherwise, when spanning reads are not available, it is recommended that the specific repeat number be reported along with confidence intervals or ranges returned by the software. If the point estimate falls in one classification (eg, intermediate or reduced penetrance), but the confidence interval overlaps with another category (eg, premutation or full penetrance), the repeat number could be reported with an appropriate recommendation for follow-up confirmation.
- It is recommended that information on controls and assay conditions be outlined by the performing laboratory. The software version must be included in the report.
- Limitations of testing should be included in the report, including genes/loci not tested.
- Some REs may contain interrupted repeat motifs whose clinical implications are currently not well established. Laboratories should indicate in their interpretation the presence of such motifs of uncertain significance.
- It is the responsibility of the laboratory to empirically determine the detection limit(s) for their assays and include them in the methods and limitations section of the clinical report.
- If using a laboratory-developed test instead of a test approved by the Food and Drug Administration, the following statement must be included in the report: "This test was developed and its performance characteristics determined by clinical laboratory following CAP guidelines (MOL.30785 and MOL.31590) or "specific" state guidelines. This test has not been cleared or approved by the US Food and Drug Administration. The laboratory is regulated under Clinical Laboratory Improvement Amendment (CLIA)'88 as qualified to perform high-complexity

testing. This test is intended for clinical purposes and should not be regarded as investigational or for research."

Clinical considerations

Diagnosis and counseling

The accurate diagnosis of REDs is essential for the application of appropriate therapies and management. Most REDs typically feature strong genotype-phenotype correlations, with the size of the expansion correlating with the age of onset and disease progression.[5] Strong clinical suspicion even in the absence of family history requires diagnostic testing. For example, *FMR1* is associated with 3 relatively distinct disorders depending on repeat size, fragile X syndrome, FXPOI (Primary Ovarian Insufficiency, Fragile X-associated, OMIM 311360), and FXTAS (Fragile X Tremor and Ataxia Syndrome, OMIM 300623), each of which affects different patient subgroups and requires unique reporting, counseling, and management considerations.[38] Each of these disorders is correlated with a different RE length and therefore different downstream effects.

Test ordering and report evaluation

Laboratory requests for REDs range from asymptomatic testing, carrier screening, prenatal diagnosis, and diagnostic testing for individuals with phenotypic findings. Reporting considerations differ in diagnostic test requests used for known family history or symptomatic cases compared with screening requests for prenatal testing or because of suspicion of family history. In general, because of the frequent limited availability of clinical information associated with requests for RED testing, laboratory reports should not include predictions of symptoms or age of onset based on expansion size. In addition, some tests may include detection of interruptions within the RE; unless there is a well-established clinical utility of reporting variant repeats with interruptions, it is not recommended to predict disease prognosis based on the detection of expanded alleles with interruptions.

Prenatal testing

Given that many REDs are associated with adult-onset neurodegenerative diseases, the prenatal screening and diagnosis of REDs require extensive genetic counseling. If prenatal testing is performed, it may not be possible to accurately predict phenotypic outcomes for some disorders. Prenatal screening may bring challenges for other family members as well, including the inadvertent diagnosis of symptomatic or asymptomatic carriers (eg, FXTAS in the father of a carrier mother). Counseling should include the consideration of follow-up testing for some members of the family as indicated, as well as the availability of prenatal diagnosis and pregnancy management options. REDs may be diagnosed in the presymptomatic setting during routine carrier screening (such as fragile X syndrome on expanded carrier screening panels), and laboratory reports should include recommendations for genetic counseling if identified.

With respect to late-onset disorders, prognosis and management may differ significantly based on the specific RED. In general, the earlier the age of onset and the larger the RE, the more severe the phenotype.[2] Given genetic anticipation, individuals with milder, adult-onset disorders may have offspring with more severe and earlier onset phenotypes.

Conclusion

Recent advances in NGS technologies coupled with the development of bioinformatic pipelines have enabled the detection of REs for a variety of REDs. Although NGS offers high-throughput capabilities and flexibility in read length, short-read platforms often struggle with accurately genotyping large REs, and long-read technologies can suffer from error rates requiring robust bioinformatic correction. Additionally, challenges such as computational complexity, somatic mosaicism detection, and lack of standardized pipelines highlight the need for necessary specifications on how best to apply NGS for RE detection. Ongoing advancements in sequencing and data analysis are expected to address these limitations.

Acknowledgments

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States Government. The findings and conclusions presented in this article are those of the author(s) and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services.

Conflict of Interest

Saurav Guha, Akanchha Kesari, Shreshtha Garg, Nancy C. Rose, and Bryce A. Seifert receive salary for providing clinical services that may be relevant to the content of this document in either the laboratory or patient care setting at their listed affiliations. Bryce A. Seifert is supported by the Division of Intramural Research/NIH. All other authors declare no conflicts of interest.

Affiliations

[1]New York Genome Center, New York, NY; [2]Department of Medical Genetics, The University of British Columbia, Vancouver, BC, Canada; [3]Children's and Women's

Hospital, Vancouver, BC, Canada; [4]Illumina, San Diego, CA; [5]Helix, Inc, San Mateo, CA; [6]Division of Maternal Fetal Medicine, Department of Obstetrics and Gynecology, University of Utah, Salt Lake City, UT; [7]Division of Intramural Research, NIAID, Bethesda, MD; [8]Belay Diagnostics, Chicago, IL; [9]American College of Medical Genetics and Genomics, Bethesda, MD

References

1. Malik I, Kelley CP, Wang ET, Todd PK. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat Rev Mol Cell Biol*. 2021;22(9):589-607. http://doi.org/10.1038/s41580-021-00382-6
2. Depienne C, Mandel JL. 30 years of repeat expansion disorders: what have we learned and what are the remaining challenges? *Am J Hum Genet*. 2021;108(5):764-785. http://doi.org/10.1016/j.ajhg.2021.03.011
3. Rajan-Babu IS, Dolzhenko E, Eberle MA, Friedman JM. Sequence composition changes in short tandem repeats: heterogeneity, detection, mechanisms and clinical implications. *Nat Rev Genet*. 2024;25(7):476-499. http://doi.org/10.1038/s41576-024-00696-z
4. Ibañez K, Jadhav B, Zanovello M, et al. Increased frequency of repeat expansion mutations across different populations. *Nat Med*. 2024;30(11):3357-3368. http://doi.org/10.1038/s41591-024-03190-5
5. Paulson H. Repeat expansion diseases. *Handb Clin Neurol*. 2018;147:105-123. http://doi.org/10.1016/B978-0-444-63233-3.00009-9
6. Wallace SE, Bean LJH. Resources for genetics professionals—genetic disorders caused by nucleotide repeat expansions and contractions. In: Adam MP, Feldman J, Mirzaa GM, et al., eds. *GeneReviews*. University of Washington; 2017. Accessed September 8, 2024. https://www.ncbi.nlm.nih.gov/books/NBK535148/
7. Hiatt L, Weisburd B, Dolzhenko E, et al. STRchive: a dynamic resource detailing population-level and locus-specific insights at tandem repeat disease loci. *Genome Med*. 2025;17(1):29. http://doi.org/10.1186/s13073-025-01454-4
8. Bean L, Bayrak-Toydemir P. American College of Medical Genetics and Genomics Standards and Guidelines for Clinical Genetics Laboratories, 2014 edition: technical standards and guidelines for Huntington disease. *Genet Med*. 2014;16(12):e2. http://doi.org/10.1038/gim.2014.146
9. Prior TW. Technical standards and guidelines for myotonic dystrophy type 1 testing. *Genet Med*. 2009;11(7):552-555. http://doi.org/10.1097/GIM.0b013e3181abce0f
10. Spector E, Behlmann A, Kronquist K, et al. Laboratory testing for fragile X, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(5):799-812. http://doi.org/10.1038/s41436-021-01115-y
11. Aydin G, Dekomien G, Hoffjan S, Gerding WM, Epplen JT, Arning L. Frequency of SCA8, SCA10, SCA12, SCA36, FXTAS and C9orf72 repeat expansions in SCA patients negative for the most common SCA subtypes. *BMC Neurol*. 2018;18(1):3. http://doi.org/10.1186/s12883-017-1009-9
12. Turro E, Astle WJ, Megy K, et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature*. 2020;583(7814):96-102. http://doi.org/10.1038/s41586-020-2434-2
13. Ashley EA. Towards precision medicine. *Nat Rev Genet*. 2016;17(9):507-522. http://doi.org/10.1038/nrg.2016.86
14. Mousavi N, Shleizer-Burko S, Yanicky R, Gymrek M. Profiling the genome-wide landscape of tandem repeat expansions. *Nucleic Acids Res*. 2019;47(15):e90. http://doi.org/10.1093/nar/gkz501
15. Tankard RM, Bennett MF, Degorski P, Delatycki MB, Lockhart PJ, Bahlo M. Detecting expansions of tandem repeats in cohorts sequenced with short-read sequencing data. *Am J Hum Genet*. 2018;103(6):858-873. http://doi.org/10.1016/j.ajhg.2018.10.015
16. Rajan-Babu IS, Peng JJ, Chiu R, et al. Genome-wide sequencing as a first-tier screening test for short tandem repeat expansions. *Genome Med*. 2021;13(1):126. http://doi.org/10.1186/s13073-021-00932-9
17. Bahlo M, Bennett MF, Degorski P, Tankard RM, Delatycki MB, Lockhart PJ. Recent advances in the detection of repeat expansions with short-read next-generation sequencing. *F1000Res*. 2018;7. F1000 Faculty Rev-736 https://doi.org/10.12688/f1000research.13980.1
18. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun*. 2021;9(1):98. http://doi.org/10.1186/s40478-021-01201-x
19. Fleszar Z, Dufke C, Sturm M, et al. Short-read genome sequencing allows "en route" diagnosis of patients with atypical Friedreich ataxia. *J Neurol*. 2023;270(8):4112-4117. http://doi.org/10.1007/s00415-023-11745-8
20. Ibañez K, Polke J, Hagelstrom RT, et al. Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study. *Lancet Neurol*. 2022;21(3):234-245. http://doi.org/10.1016/S1474-4422(21)00462-2
21. van der Sanden BPGH, Corominas J, de Groot M, et al. Systematic analysis of short tandem repeats in 38,095 exomes provides an additional diagnostic yield. *Genet Med*. 2021;23(8):1569-1573. http://doi.org/10.1038/s41436-021-01174-1
22. Seifert BA, Reddi HV, Kang BE, et al. Myotonic dystrophy type 1 testing, 2024 revision: A technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2024;26(8):101145. http://doi.org/10.1016/j.gim.2024.101145
23. Latham GJ, Coppinger J, Hadd AG, Nolin SL. The role of AGG interruptions in fragile X repeat expansions: a twenty-year perspective. *Front Genet*. 2014;5:244. http://doi.org/10.3389/fgene.2014.00244
24. Matsuyama Z, Izumi Y, Kameyama M, Kawakami H, Nakamura S. The effect of CAT trinucleotide interruptions on the age at onset of spinocerebellar ataxia type 1 (SCA1). *J Med Genet*. 1999;36(7):546-548. http://doi.org/10.1136/jmg.36.7.546
25. Braida C, Stefanatos RKA, Adam B, et al. Variant CCG and GGC repeats within the CTG expansion dramatically modify mutational dynamics and likely contribute toward unusual symptoms in some myotonic dystrophy type 1 patients. *Hum Mol Genet*. 2010;19(8):1399-1412. http://doi.org/10.1093/hmg/ddq015
26. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35(22):4754-4756. http://doi.org/10.1093/bioinformatics/btz431
27. Dolzhenko E, Weisburd B, Ibañez K, et al. REViewer: haplotype-resolved visualization of read alignments in and around tandem repeats. *Genome Med*. 2022;14(1):84. http://doi.org/10.1186/s13073-022-01085-z
28. Rajan-Babu IS, Law HY, Yoon CS, Lee CG, Chong SS. Simplified strategy for rapid first-line screening of fragile X syndrome: closed-tube triplet-primed PCR and amplicon melt peak analysis. *Expert Rev Mol Med*. 2015;17:e7. http://doi.org/10.1017/erm.2015.5
29. Halman A, Dolzhenko E, Oshlack A. STRipy: a graphical application for enhanced genotyping of pathogenic short tandem repeats in sequencing data. *Hum Mutat*. 2022;43(7):859-868. http://doi.org/10.1002/humu.24382
30. Cui Y, Ye W, Li JS, et al. A genome-wide spectrum of tandem repeat expansions in 338,963 humans. *Cell*. 2024;187(9):2336-2341.e5. http://doi.org/10.1016/j.cell.2024.03.004
31. Dolzhenko E, van Vugt JJFA, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. *Genome Res*. 2017;27(11):1895-1903. http://doi.org/10.1101/gr.225672.117
32. Rehder C, Bean LJH, Bick D, et al. Next-generation sequencing for constitutional variants in the clinical laboratory, 2021 revision: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23(8):1399-1415. http://doi.org/10.1038/s41436-021-01139-4
33. Crooks KR, Farwell Hagman KD, Mandelker D, et al. Recommendations for next-generation sequencing germline variant confirmation:

ARTICLE IN PRESS

a joint report of the Association for Molecular Pathology and National Society of Genetic Counselors. *J Mol Diagn*. 2023;25(7):411-427. http://doi.org/10.1016/j.jmoldx.2023.03.012

34. de Boer EN, Scheper AJ, Hendriksen D, et al. Nanopore long-read sequencing as a first-tier diagnostic test to detect repeat expansions in neurological disorders. *Int J Mol Sci*. 2025;26(7):2850. http://doi.org/10.3390/ijms26072850
35. van der Sanden B, Neveling K, Pang AWC, et al. Optical genome mapping for applications in repeat expansion disorders. *Curr Protoc*. 2024;4(7):e1094. http://doi.org/10.1002/cpz1.1094
36. Facchini S, Dominik N, Manini A, et al. Optical genome mapping enables detection and accurate sizing of RFC1 repeat expansions. *Biomolecules*. 2023;13(10):1546. http://doi.org/10.3390/biom13101546
37. den Dunnen JT, Dalgleish R, Maglott DR, et al. HGVS recommendations for the description of sequence variants: 2016 update. *Hum Mutat*. 2016;37(6):564-569. http://doi.org/10.1002/humu.22981
38. Walsh MB, Charen K, Shubeck L, et al. Men with an FMR1 premutation and their health education needs. *J Genet Couns*. 2021;30(4):1156-1167. http://doi.org/10.1002/jgc4.1399