Reaching further into dynamic mutations with the long-read era

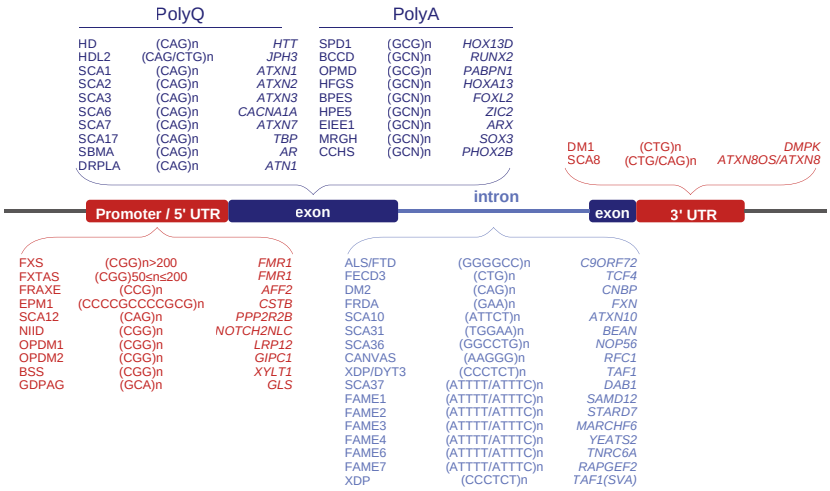
Tandem repeats

Background

Tandem repeats are a common type of genetic variation in the human genome, ranging from 2 bp to over 100 bp. More than 60 tandem repeat loci have been linked to over 69 diseases [1], including Fragile X syndrome, Huntington's disease, and spinocerebellar ataxias.

Many of these conditions exhibit high phenotypic variability and overlapping clinical features, making diagnosis based on symptoms alone unreliable and sometimes leading to incorrect molecular testing approaches. A precise, multiplex molecular assay is needed, but routine tests face technological limitations, often leaving cases unresolved.

	Southern blot	RP-PCR	NGS
Extreme GC content	✓	✓	×
Interruptive motifs	×	×	×
Large repeat units (VNTR)	✓	×	×
Number of gene per assay	1	1-10	10-30
Repeat count accuracy	Low	Low (Limited to count <200)	Low (Limited to size <150 bp)
Throughput	Low	Low	High



Challenge of routine assays:

Interruptive motifs
Not designed to detect interruptive accurately

Variability in repeat sizes (from STR to VNTR)
RP-PCR is limited to STRs (usually 2–6 bp)
NGS short-read length (~150 bp) is insufficient to span large expansion

Flanking region mutations
Off-target may cause false results in PCR assays

Extreme GC content of repeat units
Limiting the reliability of NGS because of low coverage

Most located in non-coding regions
Blind spots for whole-exome sequencing (WES)

Long-read sequencing opens new frontiers in the detection

To date, nearly half of the disorders associated with expansion loci have been discovered in the past decade, as previous methodological challenges hindered their detection.

Advances in long-read sequencing have accelerated the detection of expansions, eliminating the need for linkage data or prior hypotheses about the expanded motif.

Timeline of repeat expansion discovery in human disorders

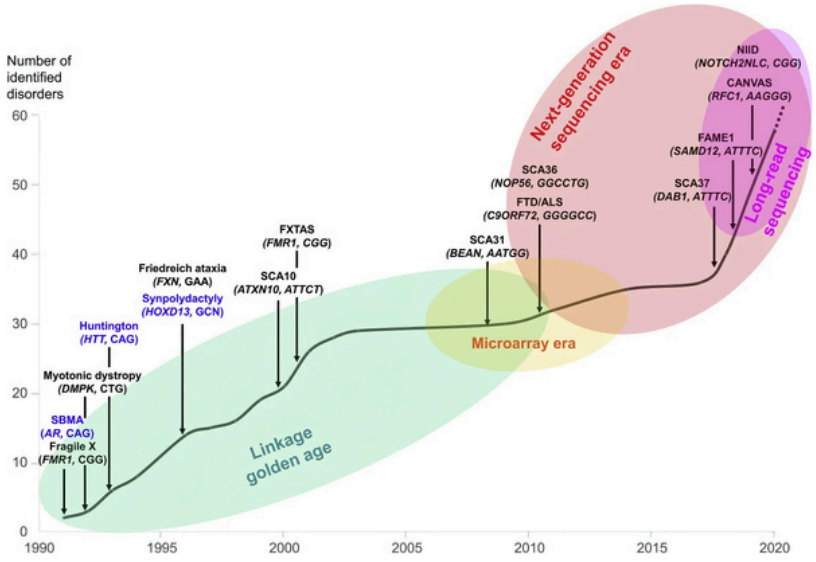


Figure adopted from reference [1] with modification

dmTGS - massive tandem repeat panel with long-read sequencing

pioneers the future of dynamic mutation detection

dmTGS overcomes these challenges using HiFi reads with SMRT sequencing. This advanced approach surpasses PCR-based assays and short-read sequencing, providing a more comprehensive and reliable genetic profile for disease diagnosis.

HiFi reads
Accuracy > 99.9%
Long-read
Uniform coverage
No GC bias

dmTGS
Repeat expansions
63 genes, 68 diseases

Technology: Single molecule real-time (SMRT) sequencing
Platform: PacBio Vega, Sequel II, and Sequel IIE system
Sample type: Blood and gDNA
Turnaround time: 29 working days
(starting from the date of sample arrival at the testing laboratory)

dmTGS methodologies

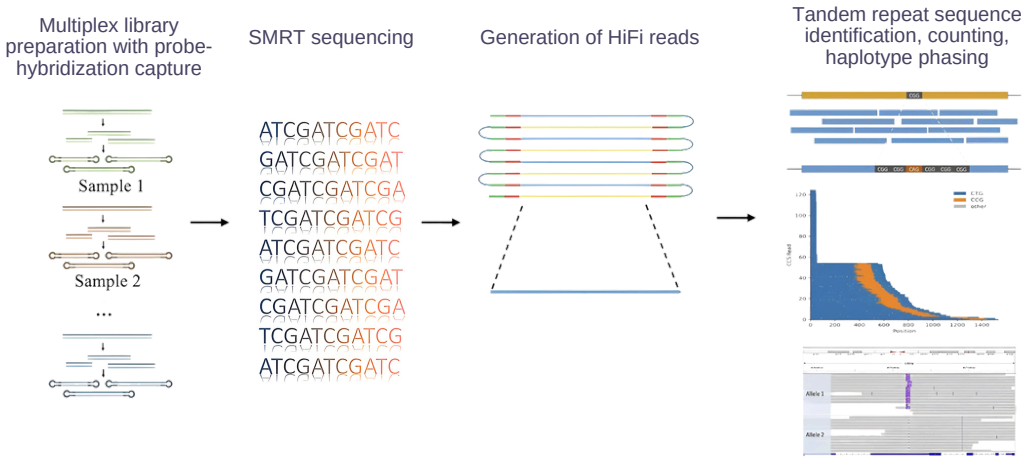


Figure adopted from reference [2] with modification

Comprehensive detection

Covers 68 diseases associated with 63 genes in a single assay, including the hard-to-map regions

High resolution

Precisely distinguishes pathogenic from non-pathogenic repeat counts at a single-repeat level, reducing uncertainty in borderline cases

Sensitive to low-level mosaicism

Detects mosaicism at low levels (e.g. 1% for *FMR1* CGG premutations and 5% for full mutations [2])

Accurate detection of interruptive motifs

Detects interruptions within repetitive sequences that can impact genetic stability, transmission, and pathogenicity
For example:

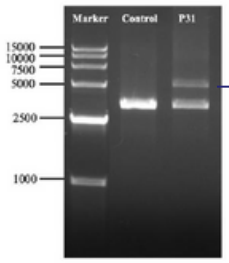
- For Fragile X syndrome, women with AGG interruptions in CGG repeats are less likely to have children with full mutations [3]
- Similar mechanisms are observed in interruptive motifs in genes such as *ATXN1*, *ATXN2*, *DMPK*, and *HTT* [4, 5]

Detailed characterization of polymorphic sequences

- Highly polymorphic sequences (e.g. VNTR) sometimes span dozens of kilobases
- Posing challenges for PCR-based assays (large repeat unit) and nanopore (high base error rates)
- dmTGS, with its single-nucleotide resolution, overcomes these challenges by enabling accurate mapping, repeat calling, and precise delineation of complex sequences and motifs

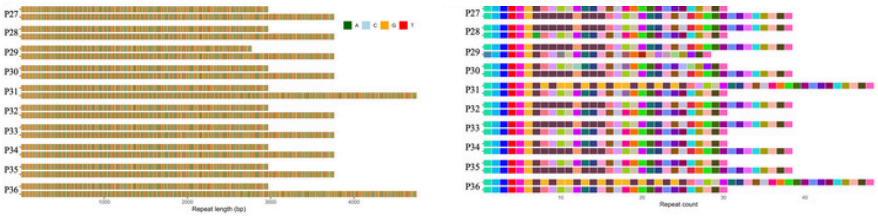
dmTGS - massive tandem repeat panel with long-read sequencing

Analysis of the 99-mer VNTR in *PLIN4* among 10 patients



PCR assays amplification can only provide an approximate estimate of the repeat number based on fragment size

In contrast, dmTGS accurately determined repeat counts in heterogeneous repeats. Sequence variability in the *PLIN4* gene is illustrated in composition plots.



In one notable patient (P53) with phenotypes highly suspected with myotonic dystrophy:

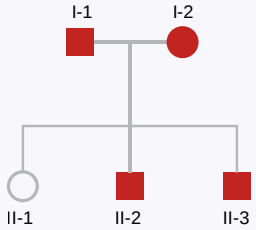
1. RP-PCR assay for the *DMPK* gene (myotonic dystrophy type I) indicated CTG repeat count of 5/13 (negative)
2. Repeat analysis using "Expansion Hunter" on WES data reported a repeat count of 13/13 (negative)
3. dmTGS analysis detected a repeat count of 13/586 (positive) and identified a large CCG interrupting motif at the 3′ end of the repeats, potentially interfering with RP-PCR extension

Information and figure derived from reference [2]

Case study

Clinical Information

- I-1 (Proband): Slurred speech, choking easily when drinking water, urinary incontinence, weakness in both lower limbs
- I-2 (Wife): Slurred speech, difficulty walking
- II-1 (Daughter): Asymptomatic
- II-2 (Elder son): Slurred speech, difficulty walking
- II-3 (Younger son): Slurred speech, difficulty walking



Previous Genetic Test Results

1. WES + CNV-seq for I-1, I-2, and II-2: Negative
2. RP-PCR panel for spinocerebellar ataxia (10 genes) for I-1: Spinocerebellar ataxia 3 (SCA3) positive
3. SCA3 test for other members:
   - II-1: Negative
   - II-2: Positive
   - II-3: Negative

The underlying cause of the phenotype in I-2 and II-3 remains unknown.

To further investigate the cause, dmTGS is performed:

- I-2 and II-3: Neuronal intranuclear inclusion disease (NIID) (*NOTCH2NLC* 22/68) positive
- II-2: SCA3 (*ATXN3* 16/73) positive and neuronal intranuclear inclusion disease (NIID) (*NOTCH2NLC* 22/69) positive
- Abnormal expansion in the *NOTCH2NLC* gene is associated with NIID, which matches the clinical phenotypes of the I-2 and II-3

In the Chinese population, SCA3 is the most common ataxia. While SCA3 or SCA10 panel testing is preferred for typical cases, overlapping phenotypes in tandem repeat disorders make symptom-based diagnosis unreliable.

dmTGS provides a comprehensive single-test solution, minimizing misdiagnosis risk.

References: [1] 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. [2] Yang K, Liu Y, Zhang J, et al. dmTGS: Precise Targeted Enrichment Long-Read Sequencing Panel for Tandem Repeat Detection. *Clin Chem*. 2025;71(2):319-331. [3] Yrigollen CM, Durbin-Johnson B, Gane L, Nelson DL, Hagerman R, Hagerman PJ, Tassone F. AGG interruptions within the maternal *FMR1* gene reduce the risk of offspring with fragile X syndrome. Genet Med. 2012 Aug;14(8):729-36. [4] Menon RP, Nethisinghe S, Faggiano S, et al. The role of interruptions in polyQ in the pathology of *SCA1*. *PLoS Genet*. 2013;9(7):e1003648. [5] Nethisinghe S, Pigazzini ML, Pemble S, et al. PolyQ Tract Toxicity in *SCA1* is Length Dependent in the Absence of CAG Repeat Interruption. *Front Cell Neurosci*. 2018;12:200.

(Global) www.xcelomglobal.com | marketing@xcelom.com | Xcelom Limited is a subsidiary of Berry Genomics responsible for international business.

dmTGS - massive tandem repeat panel with long-read sequencing

List of disorders covered by dmTGS

Ataxia (20)

Diseases	Gene	Repeat Unit
Spinocerebellar ataxia 1 (SCA1)	*ATXN1*	CAG, CAT
Spinocerebellar ataxia 2 (SCA2)	*ATXN2*	CAG, CAA
Spinocerebellar ataxia 3 (SCA3)	*ATXN3*	CAG, CAA
Spinocerebellar ataxia 4 (SCA4)	*ZFHX3*	GGC
Spinocerebellar ataxia 6 (SCA6)	*CACNA1A*	CAG
Spinocerebellar ataxia 7 (SCA7)	*ATXN7*	CAG
Spinocerebellar ataxia 8 (SCA8)	*ATXN8OS_ATXN8*	CTG·CAG
Spinocerebellar ataxia 10 (SCA10)	*ATXN10*	ATTCT
Spinocerebellar ataxia 12 (SCA12)	*PPP2R2B*	CAG
Spinocerebellar ataxia 17 (SCA17)	*TBP*	CAG, CAA
Spinocerebellar ataxia 31 (SCA31)	*BEAN1, TK2*	TGGAA
Spinocerebellar ataxia 27B, late-onset (SCA27B)	*FGF14*	GAA
Spinocerebellar ataxia 36 (SCA36)	*NOP56*	GGCCTG
Spinocerebellar ataxia 37 (SCA37)	*DAB1*	ATTTC
Spinocerebellar ataxia 51 (SCA51)	*THAP11*	CAG
Dentatorubral-pallidoluysian atrophy (DRPLA)	*ATN1*	CAG
Friedreich ataxia (FRDA)	*FXN*	GAA
Global developmental delay, progressive ataxia, and elevated glutamine (GDPAG)	*GLS*	GCA
Cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS)	*RFC1*	AAGGG
Fragile X tremor/ataxia syndrome (FXTAS)	*FMR1*	CGG, AGG

Epilepsy (8)

Diseases	Gene	Repeat Unit
Epilepsy, progressive myoclonic 1A (EPM1A)	*CSTB*	CCCCGCCCCGCG
Epilepsy, familial adult myoclonic, 1 (FAME1)	*SAMD12*	TTTCA/TTTTA
Epilepsy, familial adult myoclonic, 2 (FAME2)	*STARD7*	AAATG
Epilepsy, familial adult myoclonic, 3 (FAME3)	*MARCHF6*	TTTCA/TTTTA
Epilepsy, familial adult myoclonic, 4 (FAME4)	*YEATS2*	TTTCA
Epilepsy, familial adult myoclonic, 6 (FAME6)	*TNRC6A*	TTTCA
Epilepsy, familial adult myoclonic, 7 (FAME7)	*RAPGEF2*	TTTCA
Tremor, hereditary essential, 6 (ETM6)	*NOTCH2NLC*	GGC

Developmental disorders (8)

Diseases	Gene	Repeat Unit
Synpolydactyly 1 (SPD1)	*HOXD13*	GCN
Richieri-Costa-Pereira syndrome (RCPS)	*EIF4A3*	18-20 bp
Hand-foot-genital syndrome (HFGS)	*HOXA13*	GCN
Blepharophimosis, ptosis, and epicanthus inversus (BPES)	*FOXL2*	GCN
VACTERL association, X-linked	*ZIC3*	GCN
Oculo-auriculo-vertebral spectrum (OVAS)		
Tetralogy of Fallot (ToF)	*TBX1*	GCN
Baratela-Scott syndrome (BSS)	*XYLT1*	GGC

Other neurological disorders (18)

Diseases	Gene	Repeat Unit
Developmental and epileptic encephalopathy 1 (DEE1)	*ARX(ARX_1)*	GCN
Intellectual developmental disorder, X-linked 29 (XLID29)		
Fragile X syndrome (FXS)/ Fragile X-associated primary ovarian insufficiency (FXPOI)/ Fragile X-associated neuropsychiatric disorders (FXAND)	*FMR1*	CGG, AGG
Intellectual developmental disorder, X-linked 109 (XLID109)	*AFF2*	CCG
KINSSHIP syndrome (KINS)	*AFF3*	CGG
Autism spectrum disorder (FRA7A)	*ZNF713*	CGG
Intellectual developmental disorder, autosomal dominant, FRA12A type (MRFRA12A)	*DIP2B*	CGG
Jacobsen syndrome (JBS)	*CBL*	CCG
Huntington disease (HD)	*HTT*	CAG
Huntington disease-like 2 (HDL2)	*JPH3*	CTG
Neuropathy, hereditary sensory and autonomic, type VIII (HSAN8)	*PRDM12*	GCC
Creutzfeldt-Jakob disease (CJD)	*PRNP*	24 bp
Intellectual developmental disorder, X-linked, with isolated growth hormone deficiency	*SOX3*	GCN
Panhypopituitarism, X-linked		
Holoprosencephaly 5 (HPE5)	*ZIC2*	GCN
Developmental and epileptic encephalopathy 1 (DEE1)	*ARX(ARX_2)*	GCN
Partington syndrome (PRTS)		
Intellectual developmental disorder, X-linked 29 (XLID29)		

Neuromuscular disorders (12)

Diseases	Gene	Repeat Unit
Spinal and bulbar muscular atrophy, X-linked 1 (SBMA)	*AR*	CAG
Myotonic dystrophy 1 (DM1)	*DMPK*	CTG
Myotonic dystrophy 2 (DM2)	*CNBP*	CCTG, GCTC, TCTG
Oculopharyngodistal myopathy 3 (OPDM3)	*NOTCH2NLC*	GGC
Oculopharyngodistal myopathy 1 (OPDM1)	*LRP12*	CGG
Oculopharyngodistal myopathy 2 (OPDM2)	*GIPC1*	CGG
Oculopharyngodistal myopathy 4 (OPDM4)	*RILPL1*	CGG
Oculopharyngeal myopathy with leukoencephalopathy 1 (OPML1)	*NUTM2BAS1_LOC642361*	CCG·CGG
Oculopharyngeal muscular dystrophy (OPMD)	*PABPN1*	GCN
Neuronal intranuclear inclusion disease (NIID)	*NOTCH2NLC*	GGC
Frontotemporal dementia and/or amyotrophic lateral sclerosis 1 (FTDALS1)	*C9orf72*	GGGGCC
Myopathy with rimmed ubiquitin-positive autophagic vacuolation, autosomal dominant (MRUPAV)	*PLIN4*	99 bp

Ophthalmological disorder (1)

Diseases	Gene	Repeat Unit
Corneal dystrophy, Fuchs endothelial, 3 (FECD3)	*TCF4*	CTG

Respiratory disorder (1)

Diseases	Gene	Repeat Unit
Central hypoventilation syndrome, congenital, 1, with or without Hirschsprung disease (CCHS1)	*PHOX2B*	GCN

For details, please contact **marketing@xcelom.com**

INT_TGSdmTGS2v2.02012026