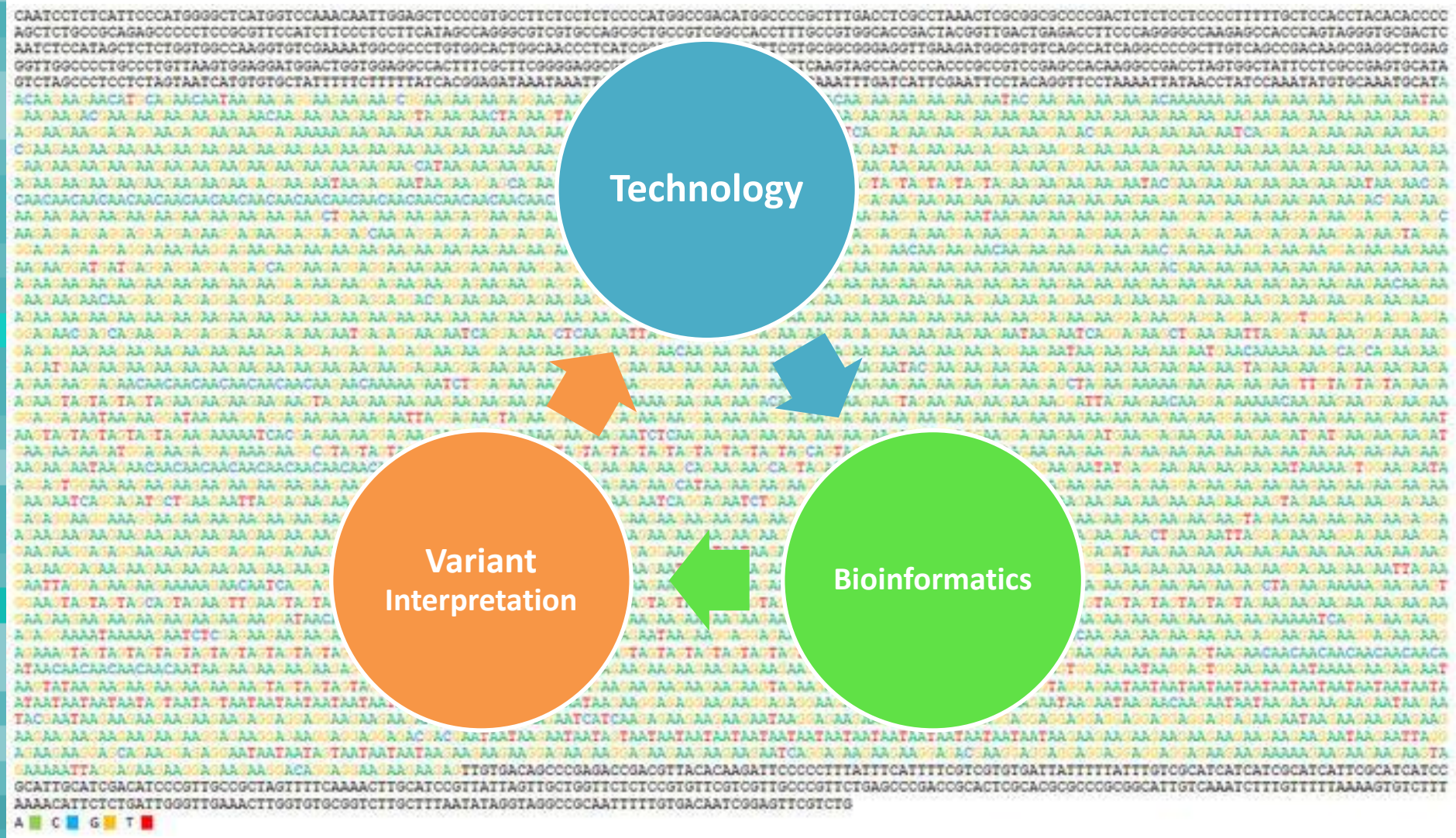
benet-pages@mgz-muenchen.de

Future Developments

Anna Benet-Pagès

Prague 08.11.2017


Technology
Variant
Interpretation
Bioinformatics

New generations?

2nd generation sequencing : short-reads

Illumina
NovaSeq

3rd generation sequencing : long-reads

Oxford nanopore

MinION
GridION
PromethION
SmidgION

3rd generation mapping

Bionano genomics

Pacbio

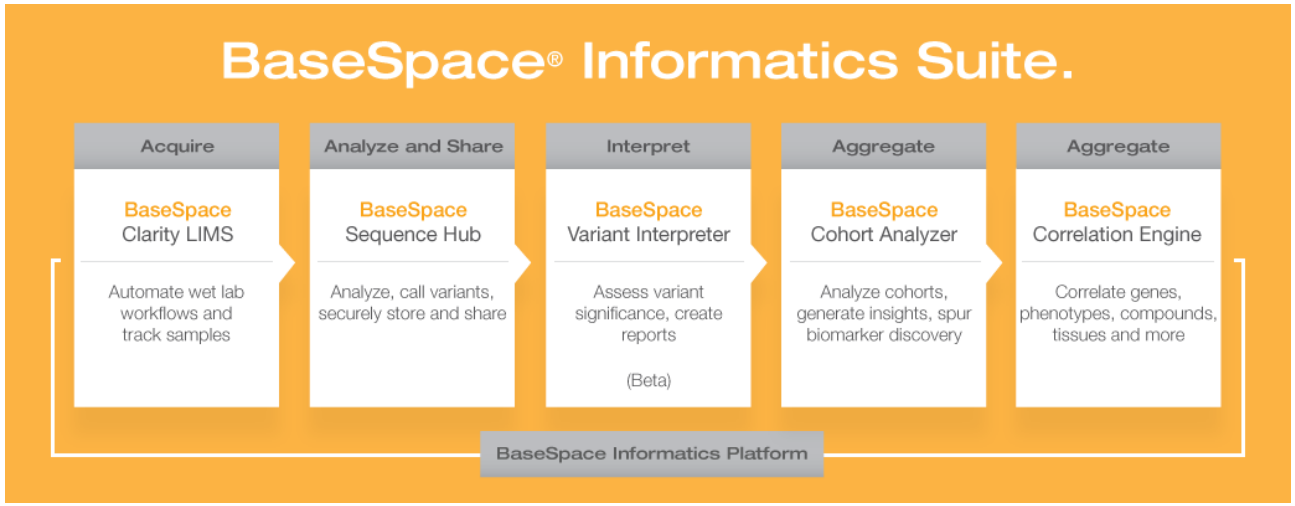
Sequel System
PacBio RS II

Short read sequencing : Illumina – NovaSeq

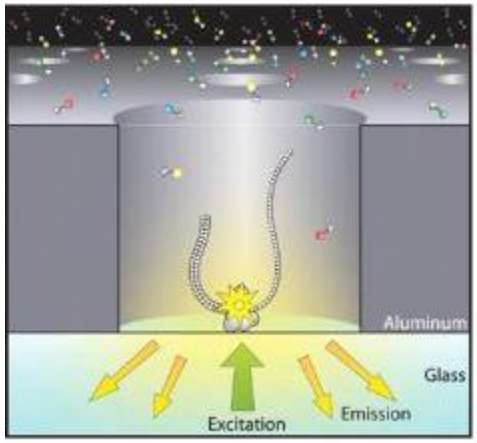
Estimated Sample Throughput for Key Applications†††

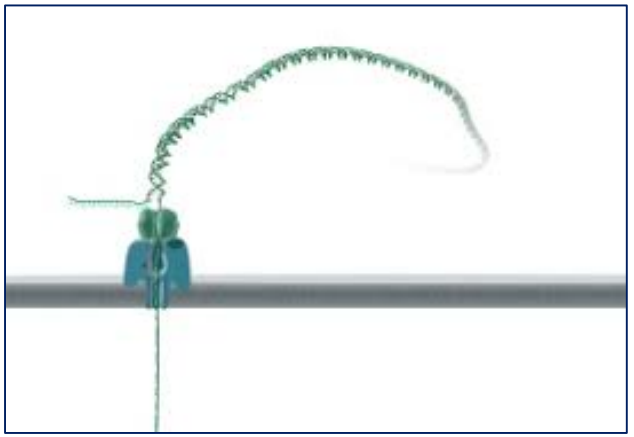
	NovaSeq 6000 System		
Flow Cell Type	S1*	S2	S4
Human Genomes per Run	~8	~16	~48
Exomes per Run	~80	~160	
Transcriptomes per Run	~64	~128	

Run Time			
2 × 50 bp	~13 hr	~16 hr	N/A
2 × 100 bp	~18 hr	~25 hr	N/A
2 × 150 bp	~24 hr	~36 hr	~44 hr

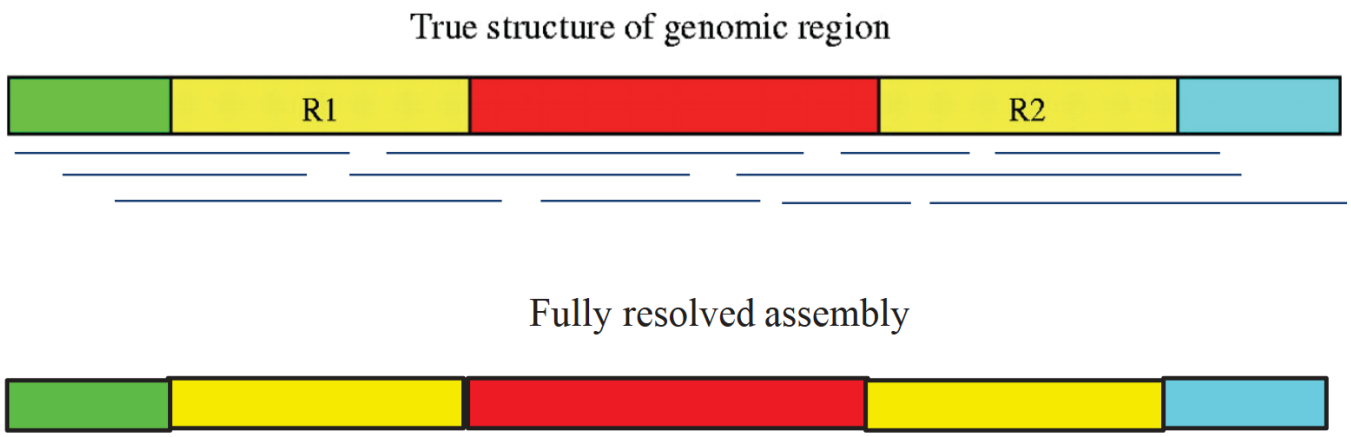


Long read sequencing : PacBio – Oxford nanopore





- Long read can span over repeated region



Long-read sequencing goes diagnostics

Long-Read Sequencing Data Enables Structural Variant Discovery For Clinical And Disease Research

Contributed Commentary By Luke Hickey

September 11, 2017 | In recent years, scientists have become increasingly aware of the clinical implication of structural variants in the human genome. These larger variants, typically defined as >50 bp, are known to cause many genetic conditions, including fragile X syndrome, Duchenne muscular dystrophy, ALS, and Tay-Sachs disease. However, due to their size, structural variants have not been as easy to study as single-nucleotide variants using short-read sequencing methods. Whether they are insertion variants that are too lengthy to span with such methods, repetitive regions that create mapping ambiguity for short reads, or GC-rich regions that challenge technologies with a systematic coverage bias, these important genomic elements have been missed by most efforts to sequence the human genome.

The advent of single molecule, real-time sequencing, with average read lengths exceeding 12 kb, has allowed scientists to reliably detect structural variation for the first time. These latest studies have shown that previous human genome data sets have dramatically underrepresented the number of structural variants—along with their effect on health and disease.

Genet Med. 2017 Jun 22. doi: 10.1038/gim.2017.86. [Epub ahead of print]

Long-read genome sequencing identifies causal structural variation in a Mendelian disease.

Merker JD[1,2], Wenger AM[3], Sneddon T[2], Grove M[2], Zappala Z[1,4], Fresard L[1], Waggott D[5,6], Utiramerur S[2], Hou Y[1], Smith KS[1], Montgomery SB[1,4], Wheeler M[5,6], Buchan JG[1,2], Lambert CC[3], Eng KS[3], Hickey L[3], Korlach J[3], Ford J[4,5,7], Ashley EA[2,4,5,6].

Long-read sequencing goes diagnostics

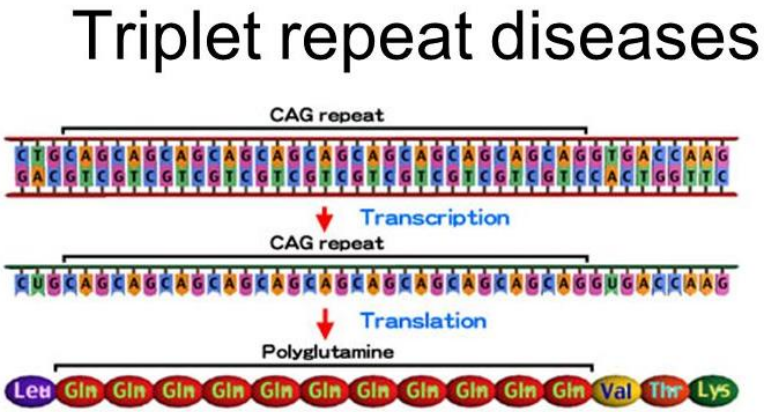


Table 2 | **Molecular features of unstable repeat expansion disorders**

Disease	Mutation/ repeat unit	Gene name (protein product)	Putative function	Normal repeat length	Pathogenic repeat length
Diseases that are caused by loss of protein function					
FRDA	$(GAA)_n$	*FRDA* (frataxin)	Mitochondrial iron metabolism	6–32	200–1,700
FRAXA	$(CGG)_n$	*FMR1* (FMRP)	Translational regulation	6–60	>200 (full mutation)
FRAXE	$(CCG)_n$	*FMR2* (FMR2)	Transcription?	4–39	200–900
Diseases that are caused by altered protein function					
SCA1	$(CAG)_n$	*SCA1* (ataxin 1)	Transcription	6–39	40–82
SCA2	$(CAG)_n$	*SCA2* (ataxin 2)	RNA metabolism	15–24	32–200
SCA3 (MJD)	$(CAG)_n$	*SCA3* (ataxin 3)	De-ubiquitylating activity	13–36	61–84
SCA6	$(CAG)_n$	*CACNA1A* ($CACNA1_A$)	P/Q-type α1A calcium channel subunit	4–20	20–29
SCA7	$(CAG)_n$	*SCA7* (ataxin 7)	Transcription	4–35	37–306
SCA17	$(CAG)_n$	*SCA17* (TBP)	Transcription	25–42	47–63
DRPLA	$(CAG)_n$	*DRPLA* (atrophin 1)	Transcription	7–34	49–88
SBMA	$(CAG)_n$	*AR* (androgen receptor)	Steroid-hormone receptor	9–36	38–62
HD	$(CAG)_n$	*HD* (huntingtin)	Signalling, transport, transcription	11–34	40–121
Diseases that are caused by altered RNA function					
DM1	$(CTG)_n$	*DMPK* (DMPK)	RNA-mediated	5–37	50–1,000
DM2	$(CCTG)_n$	*ZNF9* (ZNF9)	RNA-mediated	10–26	75–11,000
FXTAS	$(CGG)_n$	*FMR1* (FMRP)	RNA-mediated	6–60	60–200 (premutation)
Diseases of unknown pathogenic mechanism(s)					
SCA8	$(CTG)_n$	*SCA8* (transcribed/ untranslated)	Unknown	16–34	>74
SCA10	$(ATTCT)_n$	Unknown	Unknown	10–20	500–4,500
SCA12	$(CAG)_n$	*PPP2R2B* (PPP2R2B)	Phosphatase regulation	7–45	55–78
HDL2	$(CTG)_n$	*JPH3* (junctophilin 3)	PM/ER junction protein	7–28	66–78

CACNA1A, calcium channel, voltage-dependent, P/Q type, α1A subunit; *DMPK*, dystrophia myotonica protein kinase; DRPLA, dentatorubral-pallidoluysian atrophy; DM, dystrophia myotonica; ER, endoplasmic reticulum; FRDA, Friedreich ataxia; FMRP, fragile X mental retardation protein; *FMR1*, fragile X mental retardation 1; *FMR2*, fragile X mental retardation 2; FXTAS, fragile X tremor/ataxia syndrome; HD, Huntington disease; HDL2, Huntington disease-like 2; MJD, Machado–Joseph disease; PM, plasma membrane; *PPP2R2B*, protein phosphatase 2 (formerly 2A) regulatory subunit B; SBMA, spinal and bulbar muscular atrophy; SCA, spinocerebellar ataxia; TBP, TATA box binding protein; *ZNF9*, zinc-finger protein 9.

Long-read sequencing goes diagnostics

Triplet repeat diseases

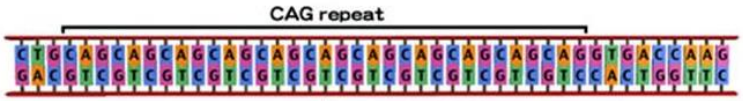


Table 2 | **Molecular features of unstable repeat expansion disorders**

Disease	Mutation/ repeat unit	Gene name (protein product)	Putative function	Normal repeat length	Pathogenic repeat length
Diseases that are caused by loss of protein function					
FRDA	$(GAA)_n$	*FRDA* (frataxin)	Mitochondrial iron metabolism	6–32	200–1,700
FRAXA	$(CGG)_n$	*FMR1* (FMRP)	Translational regulation	6–60	>200 (full mutation)
FRAXE	$(CCG)_n$	*FMR2* (FMR2)	Transcription?	4–39	200–900
Diseases that are caused by altered protein function					
DRPLA	$(CAG)_n$	*DRPLA* (atrophin 1)	Transcription	7–34	49–88
Diseases of unknown pathogenic mechanism(s)					

NPJ Parkinsons Dis. 2017 Sep 5;3:27. doi: 10.1038/s41531-017-0029-x. eCollection 2017.

Parkinson's disease associated with pure ATXN10 repeat expansion.

Schüle B[#1], McFarland KN[#2], Lee K[1], Tsai YC[3], Nguyen KD[4], Sun C[4], Liu M[4], Byrne C[1], Gopi R[5], Huang N[6], Langston JW[1], Clark T[3], Gil FJJ[7], Ashizawa T[8].

Am J Hum Genet. 2017 Nov 2;101(5):700-715. doi: 10.1016/j.ajhg.2017.09.013.

Profiling of Short-Tandem-Repeat Disease Alleles in 12,632 Human Whole Genomes.

Tang H[1], Kirkness EF[2], Lippert C[1], Biggs WH[2], Fabani M[2], Guzman E[2], Ramakrishnan S[1], Lavrenko V[1], Kakaradov B[2], Hou C[2], Hicks B[1], Heckerman D[1], Och FJ[1], Caskey CT[3], Venter JC[4], Telenti A[5].

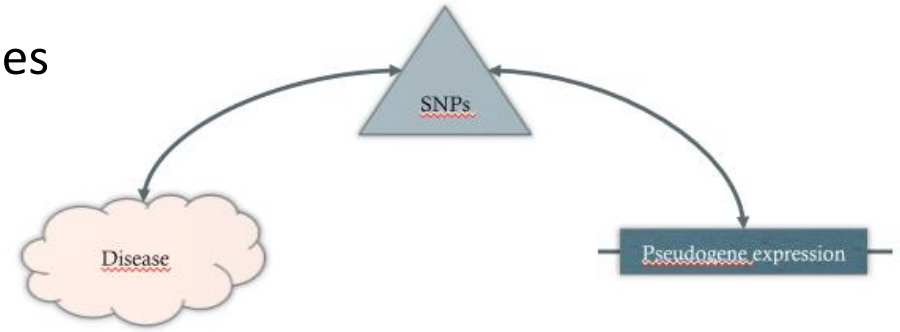
Hum Mutat. 2017 Mar;38(3):324-331. doi: 10.1002/humu.23150. Epub 2017 Jan 17.

Detecting AGG Interruptions in Male and Female FMR1 Premutation Carriers by Single-Molecule Sequencing.

Ardui S[1], Race V[1], Zablotskaya A[1], Hestand MS[1], Van Esch H[1], Devriendt K[1], Matthijs G[1], Vermeesch JR[1].

Long-read sequencing goes diagnostics

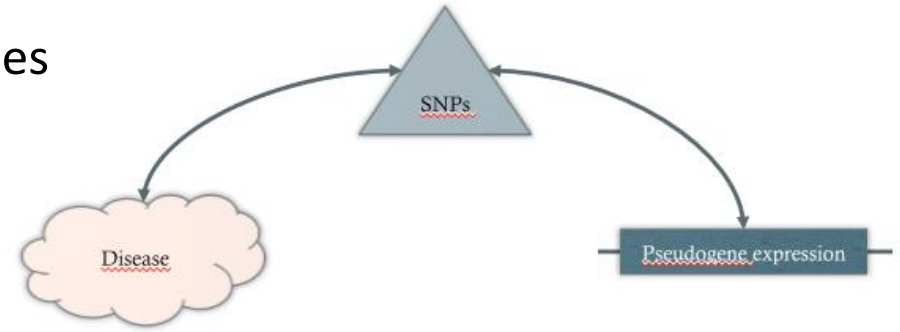
Pseudogenes / homolog / paralog genes



Pseudogene	Parental gene	Other genes	Shared microRNAs		
Oncosuppressive pseudogenes					
PTENP1	*PTEN*		miR-17, 19, 21, 26, and	Hepatocellular carcinoma Gastric cancer	(30) (45) (40)
PTENP1		*HRASLS5*	miR-135b	Breast cancer	(13)
TUSC2P	*TUSC2*		miR-17, 93, 299-3p, 520a, 608, and 661	Breast cancer	(46)
INTS6P1	*INTS6*		miR-17-5p	Hepatocellular carcinoma	(47)
Oncogenic pseudogenes					
OCT4-pg4	*OCT4*		miR-145	Hepatocellular carcinoma	(34)
OCT4-pg5	*OCT4*		miR-145	Endometrial carcinoma	(48)
HMGA1P6	*HMGA1*		miR-15, 16, 214, and 761	Thyroid carcinoma	(49)
HMGA1P7				Pituitary tumors	(50)
CYP4Z2P	*CYP4Z1*		miR-125a-3p, 197, 204, 211, and 1226	Breast cancer	(51)
BRAFP1	*BRAF*		miR-30a, 182, 590, and 876	DLBCL	(52)
Braf-rs1	*Braf*		miR-134, 543, and 653	Diffuse large B-cell lymphoma	(52)

Long-read sequencing goes diagnostics

Pseudogenes / homolog / paralog genes



Pseudogene	Parental gene	Other genes	Shared microRNAs		
Oncosuppressive pseudogenes					
PTENP1	*PTEN*		miR-17, 19, 21, 26, and	Hepatocellular carcinoma	(30)
				Gastric cancer	(45)
					(40)
PTENP1		*HRASLS5*	miR-135b	Breast cancer	(13)
BRAFP1	*BRAF*		miR-30a, 182, 590, and 876	DLBCL	(52)
Braf-rs1	*Braf*		miR-134, 543, and 653	Diffuse large B-cell lymphoma	(52)

Sci Rep. 2017 Nov 1;7(1):14789. doi: 10.1038/s41598-017-13712-6.

Nanopore-based single molecule sequencing of the D4Z4 array responsible for facioscapulohumeral muscular dystrophy.

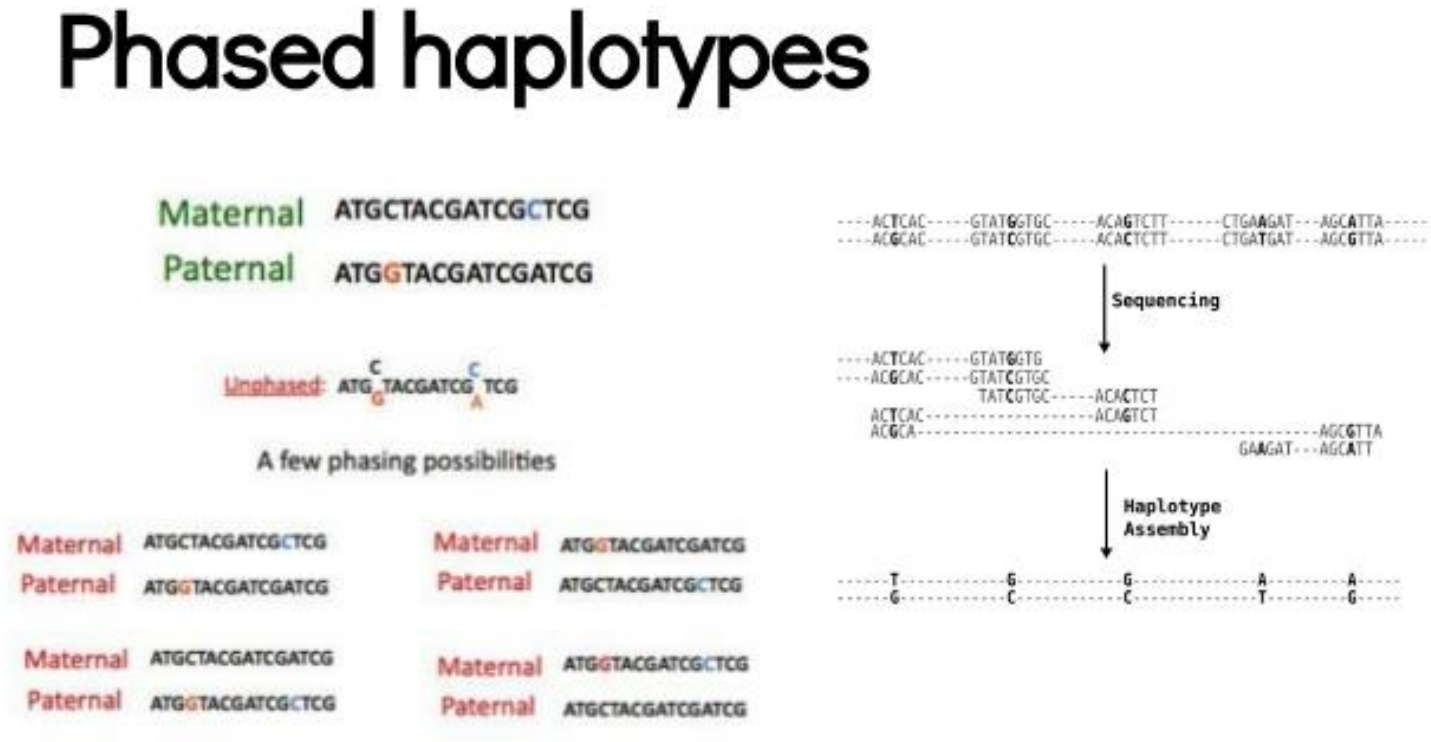
Mitsuhashi S[1,2], Nakagawa S[3,4], Takahashi Ueda M[4], Imanishi T[3], Frith MC[5,6,7], Mitsuhashi H[8].

Hum Mutat. 2017 Jul;38(7):870-879. doi: 10.1002/humu.23223. Epub 2017 May 29.

Detecting PKD1 variants in polycystic kidney disease patients by single-molecule long-read sequencing.

Borràs DM[1,2,3], Vossen RHAM[4], Liem M[4], Buermans HPJ[4], Dauwerse H[5], van Heusden D[5], Gansevoort RT[6], den Dunnen JT[4,5,7], Janssen B[1], Peters DJM[5], Losekoot M[7], Anvar SY[4,5].

Long-read sequencing goes diagnostics



Long-read sequencing goes diagnostics

HLA. 2017 Aug;90(2):79-87. doi: 10.1111/tan.13057. Epub 2017 May 25.

Dual redundant sequencing strategy: Full-length gene characterisation of 1056 novel and confirmatory HLA alleles.

Albrecht V[1], Zweiniger C[1], Surendranath V[1], Lang K[1], Schöfl G[1], Dahl A[2], Winkler S[3], Lange V[1], Böhme I[1], Schmidt AH[1,4].

Maternal ATGCTACGATCGATCG

Revealing Complete Complex KIR Haplotypes Phased By Long-Read Sequencing Technology

David Roe, Cynthia Vierra-Green, Chul-Woo Pyo, Kevin Eng, Richard Hall, Rui Kuang, Stephen Spellman, Swati Ranade, Daniel Geraghty, Martin Maiers

doi: https://doi.org/10.1101/135426

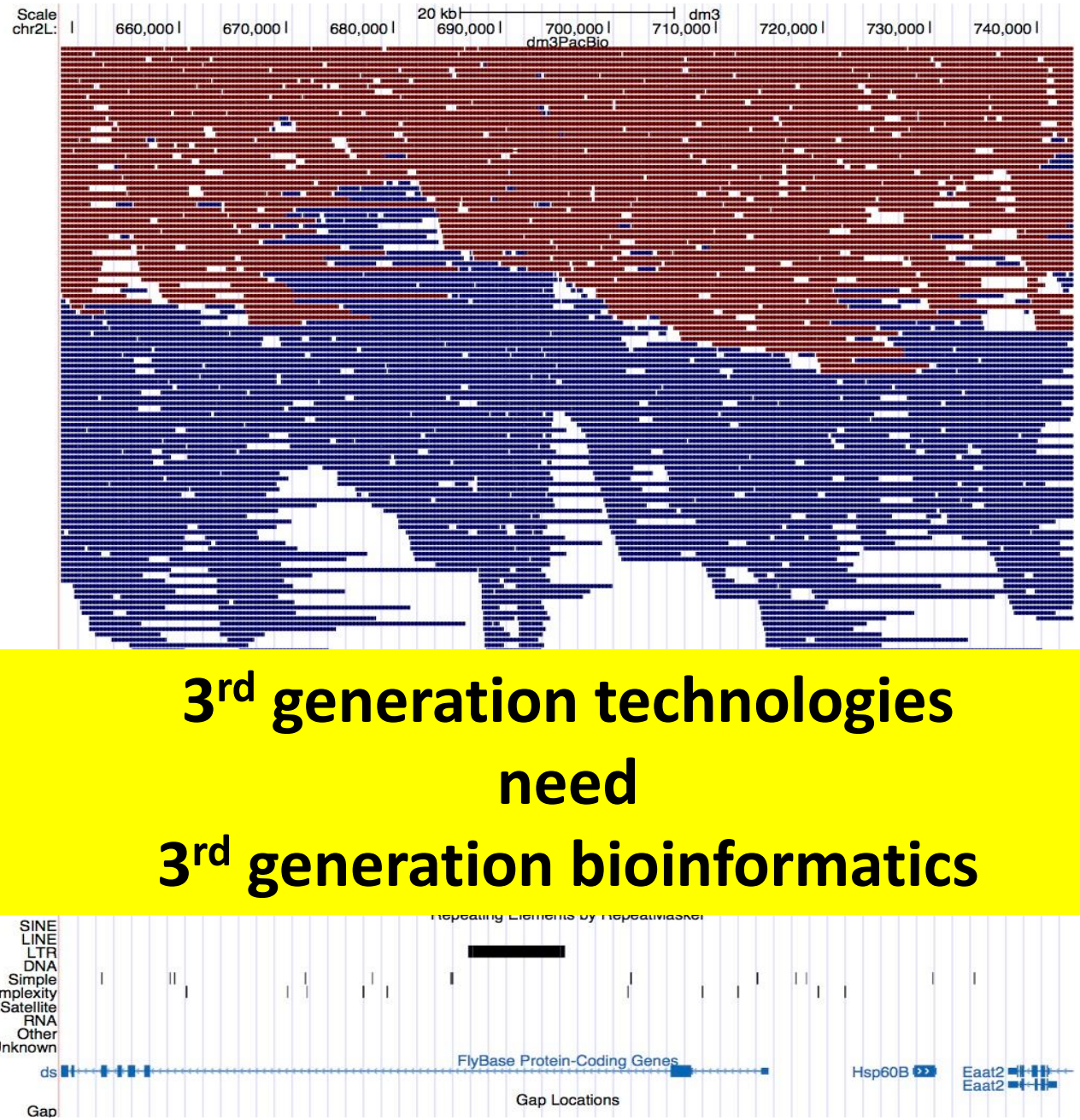
Oxoford nanopore even works in the jungle

PLoS One. 2017 Oct 4;12(10):e0184741. doi: 10.1371/journal.pone.0184741. eCollection 2017.

On site DNA barcoding by nanopore sequencing.

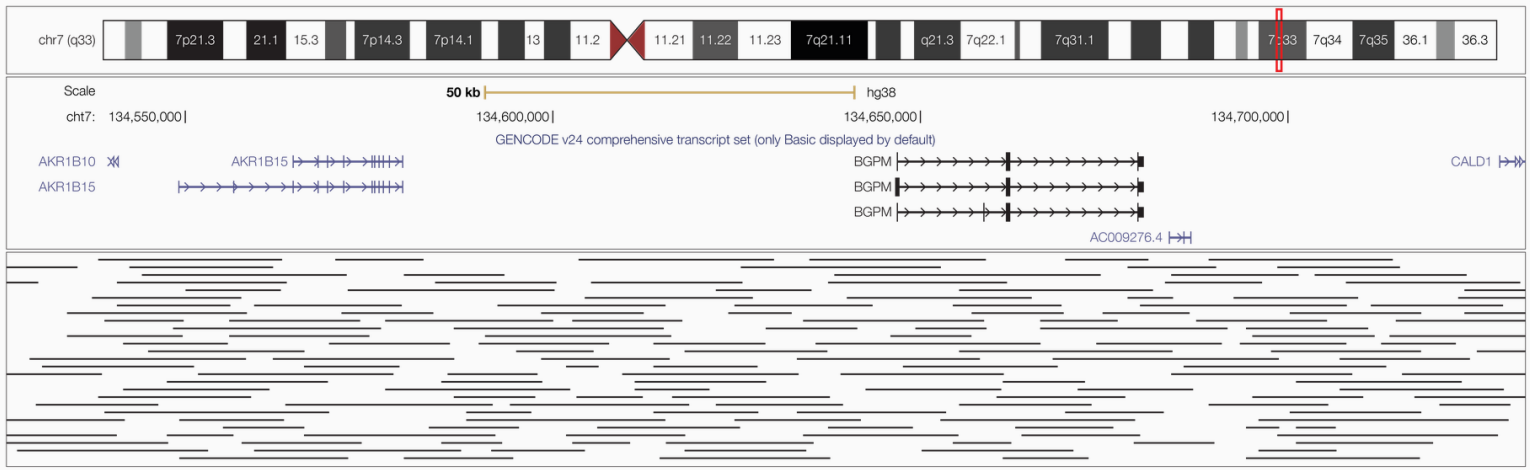
Menegon M[1], Cantaloni C[2,3], Rodriguez-Prieto A[1], Centomo C[3], Abdelfattah A[2], Rossato M[3], Bernardi M[1], Xumerle L[2,3], Loader S[4,5], Delledonne M[2,3].

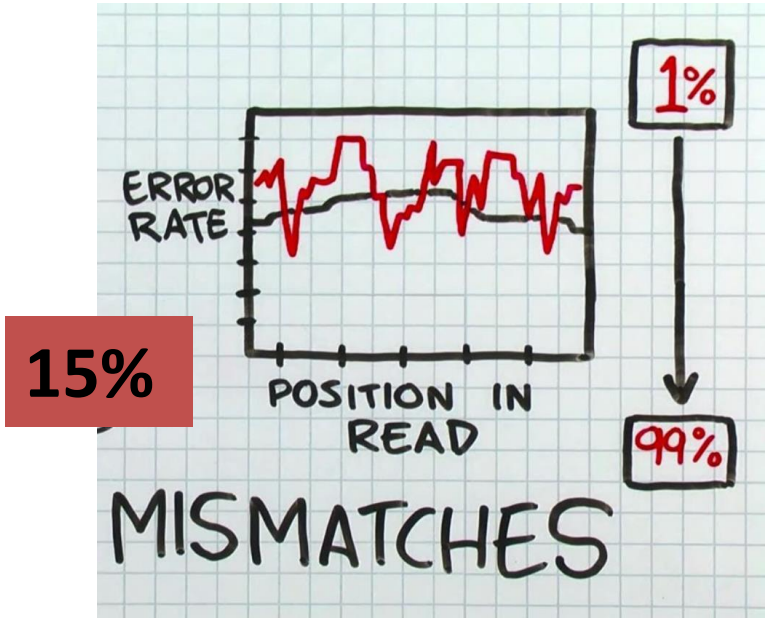
Sample species	Gene	Total Reads	2D Reads	2D Pass Reads	Similarity %			Reference (Accession number)
					Loman's	ONtoBAR	Sanger	
Amietophrynus brauni	16S	51,273	8,555	2,660	*99%*	*100%*	*100%*	*Bufo brauni (AF220886)*
Leptopelis vermiculatus	16S	109,047	57,110	42,102	*92%*	*98%*	*98%*	*Leptopelis sp. (A168408)*
Leptopelis vermiculatus	CO1	181,123	113,663	110,921	*86%*	*83%*	*82%*	*S. araneus (JF499348)*
Rieppeleon brachyurus	16S	97,080	16,760	8,026	*92%*	*100%*	*100%*	*R. brachyurus* (voucher AG19033)
Sorex alpinus	16S	84,913	24,807	7,706	*98%*	99%	*99%*	*S. alpinus (DQ630322)*
Rhynchocyon udzungwensis	CO1	167,466	104,419	97,725	88%	99%	97%	*R. petersi* (AG19033)
Arthroleptis xenodactyloides	16S	5,039	187	2	97%	100%	100%	*A. xenodactyloides* (A137057)



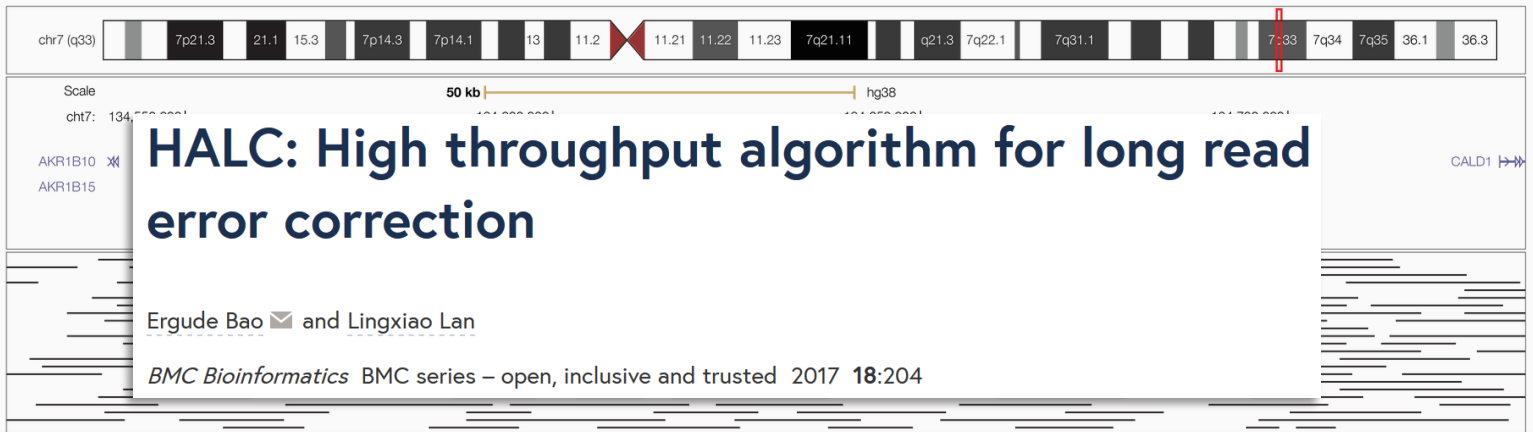
3rd generation technologies need 3rd generation bioinformatics

New bioinformatics tools for long reads





New bioinformatics tools for long reads



HALC: High throughput algorithm for long read error correction

Ergude Bao and Lingxiao Lan

BMC Bioinformatics BMC series – open, inclusive and trusted 2017 **18**:204

Bioinformatics. 2017 Aug 2. doi: 10.1093/bioinformatics/btx489. [Epub ahead of print]

LRCstats, a tool for evaluating long reads correction methods.

La S[1], Haghshenas E[2], Chauve C[1].

IEEE Trans Nanobioscience. 2017 Mar;16(2):108-115. doi: 10.1109/TNB.2017.2675981. Epub 2017 Mar 17.

HapIso: An Accurate Method for the Haplotype- Specific Isoforms Reconstruction From Long Single-Molecule Reads.

Mangul S, Yang TH, Hormozdiari F, Dainis AM, Tseng E, Ashley EA, Zelikovsky A, Eskin E.

Worth to mention…

Nucleic Acids Res. 2017 Jun 2;45(10):e76. doi: 10.1093/nar/gkx024.

MosaicHunter: accurate detection of postzygotic single-nucleotide mosaicism through next-generation sequencing of unpaired, trio, and paired samples.

Huang AY[1,2], Zhang Z[1,3], Ye AY[1,4,5], Dou Y[1,2], Yan L[1], Yang X[1], Zhang Y[6], Wei L[1].

PLEASE SELECT

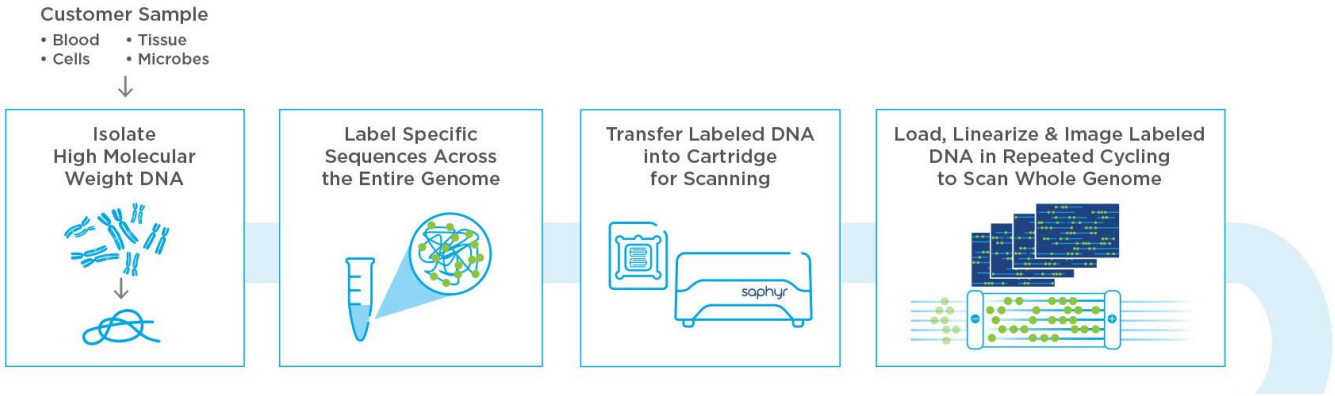
DRY LAB

BUNDLE

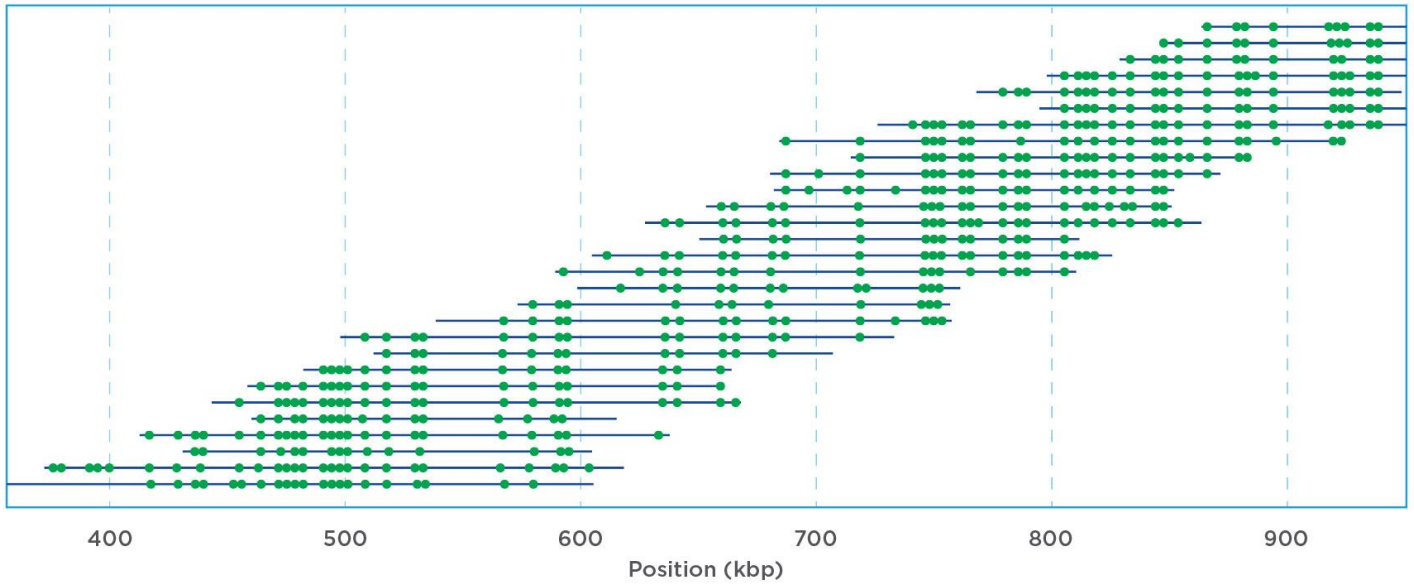
FULL SERVICE


3rd generation genome mapping

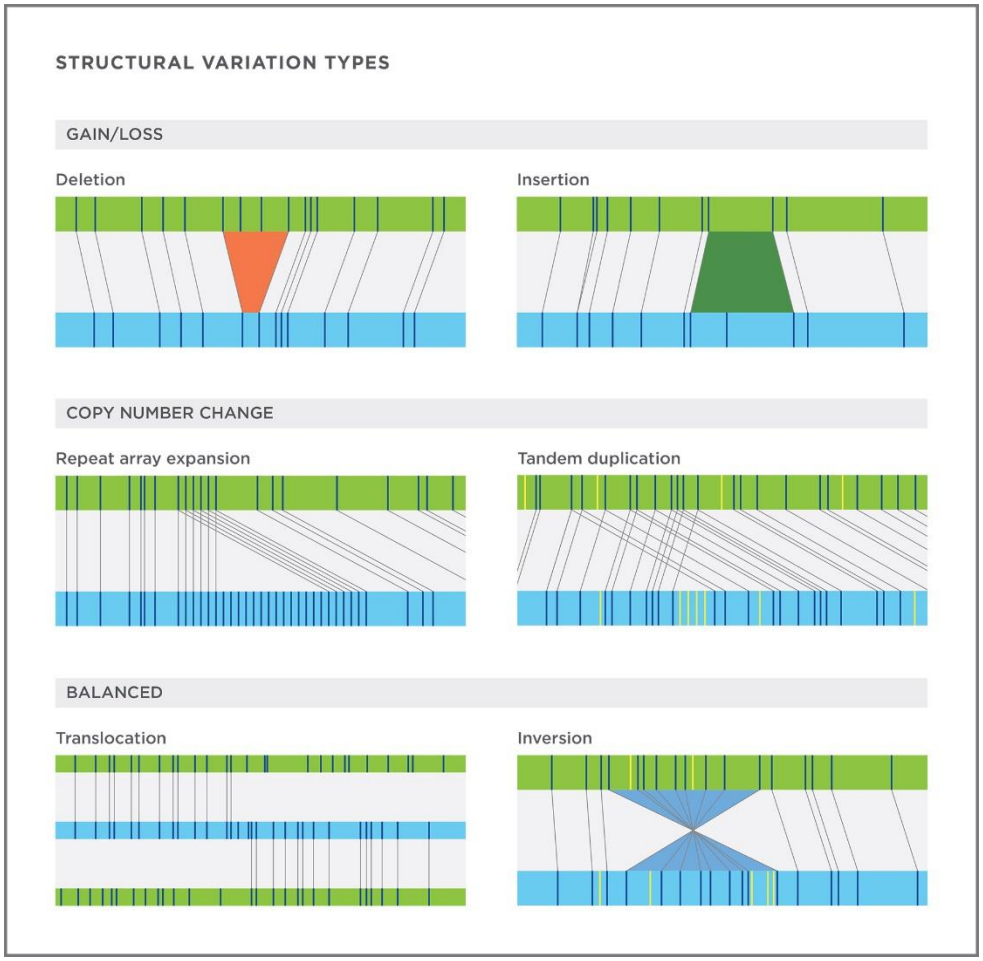
Bionano Genomics



DIGITAL REPRESENTATION OF LABELED LONG DNA

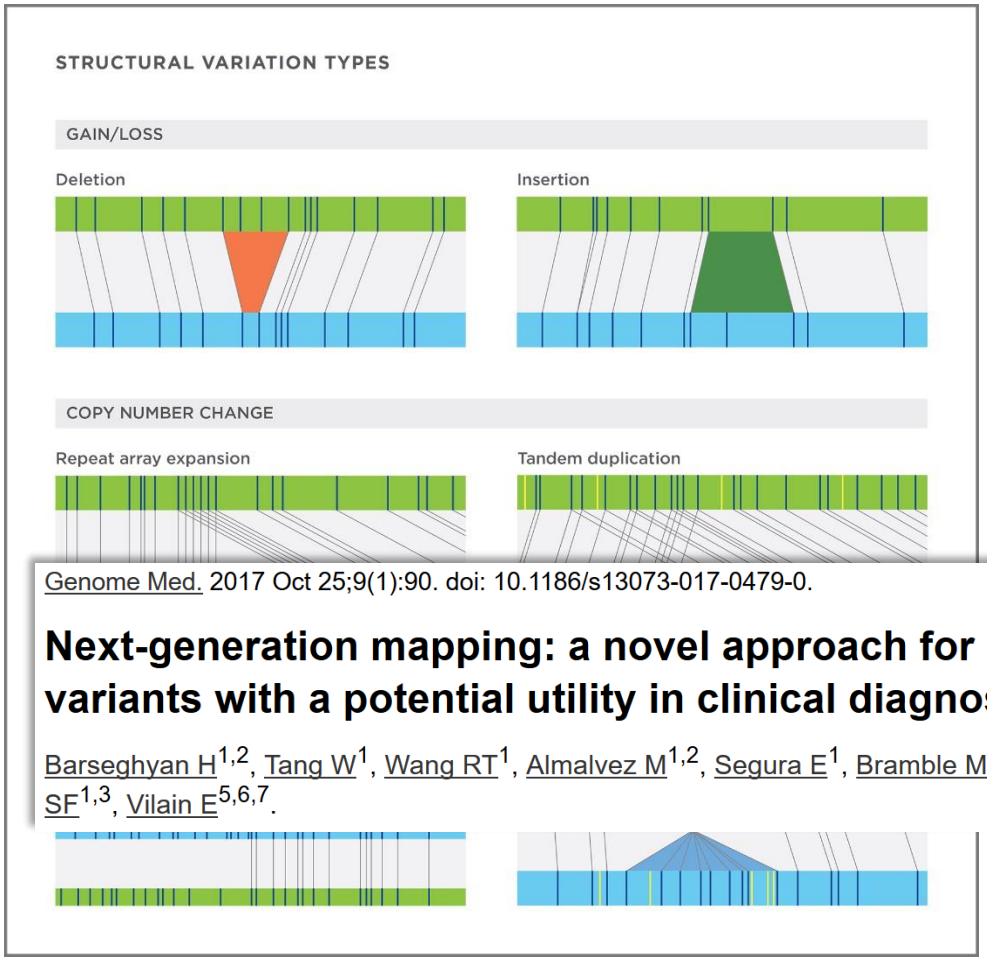


Bionano Genomics : structural variation



- 99% sensitivity for large homozygous insertions and deletions
- 87% sensitivity for large heterozygous insertions and deletions
- 98% sensitivity for translocations
- 98% sensitivity for inversions

Bionano Genomics : structural variation



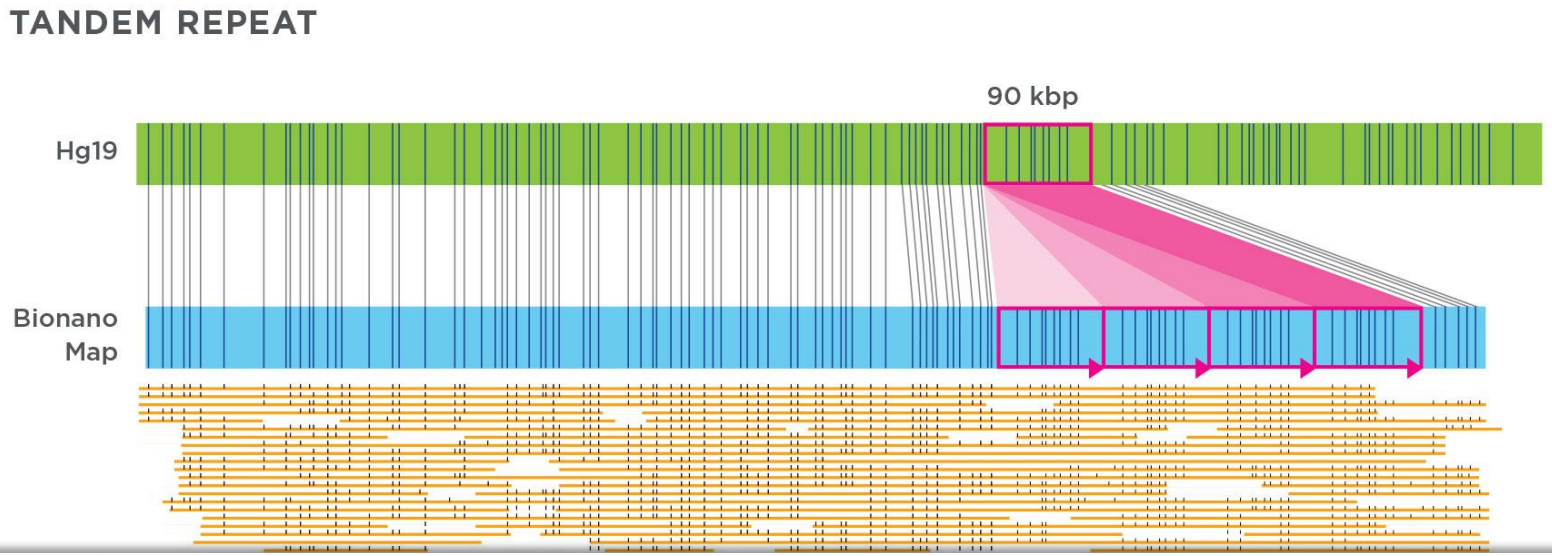
- 99% sensitivity for large homozygous insertions and deletions
- 87% sensitivity for large heterozygous insertions and deletions
- 98% sensitivity for translocations
- 98% sensitivity for inversions

Genome Med. 2017 Oct 25;9(1):90. doi: 10.1186/s13073-017-0479-0.

Next-generation mapping: a novel approach for detection of pathogenic structural variants with a potential utility in clinical diagnosis.

Barseghyan H[1,2], Tang W[1], Wang RT[1], Almalvez M[1,2], Segura E[1], Bramble MS[1,2], Lipson A[1], Douine ED[1], Lee H[3], Délot EC[1,4,2], Nelson SF[1,3], Vilain E[5,6,7].

Bionano Genomics：facioscapulohumeral muscular dystrophy



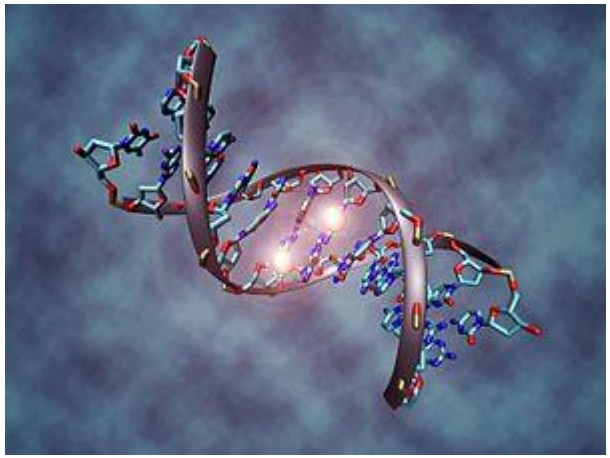
bionano GENOMICS 组 GrandOmics 希望组

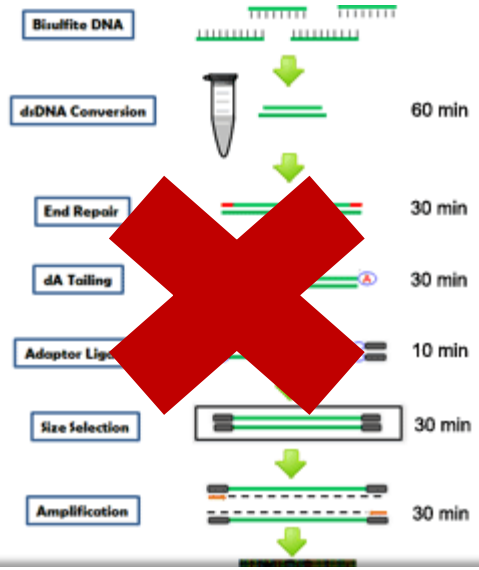
Beijing Grandomics Announces Intent to Use the Saphyr System for Molecular Diagnosis of Facioscapulohumeral Muscular Dystrophy (FSHD) in China

Beijing Grandomics using Bionano Saphyr to find Complex Genome Variants in Clincial Applications

Bionano Genomics : what else?

- Targeted methylation analysis with no needs of bisulfite conversion





New Results

Reduced representation optical methylation mapping (R^2OM^2).

Assaf Grunwald, Hila Sharim, Tslil Gabrieli, Yael Michaeli, Dmitry Torchinsky, Matyas Juhasz, Kathryn R Wagner, Jonathan Pevsner, Jeff Reifenberger, Alex R Hastie, Han Cao, Elmar Weinhold, Yuval Ebenstein

doi: https://doi.org/10.1101/113522


That's NOT all Folks!

Liquid biopsy re-emerge old technolgies

- Analysis of ctDNA in liquid biopsies requires a sensitivity down to 0.1 – 0.2%

French Firm Stilla Technologies Unveils Three-Color 'Crystal Digital' PCR Platform

Mar 16, 2016 | Madeleine Johnson

Premium

FORMULATRIX® Announces Unveiling of New High-throughput Digital PCR System

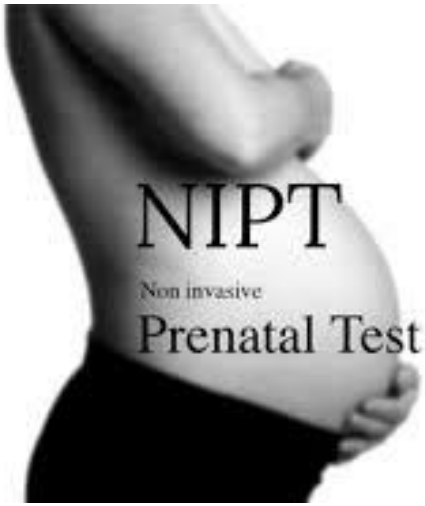
July 19, 2017 by Thomas Rawlins — Leave a Comment

BOSTON, MASSACHUSETTS—July 19, 2016— FORMULATRIX® announced today that the next generation CONSTELLATION® Digital PCR System will be revealed at the qPCR and Digital PCR Congress in Philadelphia on July 25th, 2017. "Based on customer feedback, we have built on our existing digital PCR platform by developing plates with up to 36,000 partitions per sample and by integrating partitioning, thermal cycling, and imaging into a single fully automated instrument that takes users from sample to answer in under one and a half hours," explained Jeremy Stevenson, Founder, and CEO of FORMULATRIX. "With up to five channel filters, the CONSTELLATION is the first ever digital PCR system to offer five color multiplexing making it a high-throughput platform capable of displacing qPCR as the method of choice for high-throughput quantification of nucleic acid targets."

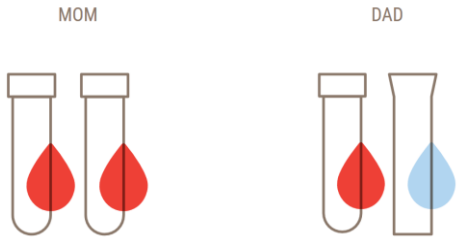
Liquid biopsy and NIPT for monogenic disorders

Baylor Genetics launches PreSeek™ - 1st non-invasive prenatal multi-gene sequencing screen

Baylor College of Medicine › Baylor College of Medicine News › Genetics › Baylor Genetics launches PreSeek™ - 1st non-invasive prenatal multi-gene sequencing screen



PreSeek™ screens for various clinically significant and life-altering genetic disorders that are not screened for with current NIPT technology. Disorders screened by this innovative test often occur in the absence of a family history of the condition. The screen, developed by the genomic experts at Baylor Genetics in conjunction with Baylor College of Medicine, assesses fetal DNA for pathogenic and likely pathogenic variants in 30 genes. PreSeek™ is the next step in the evolution of screening for genetic disorders during pregnancy, providing information that can affect medical decisions, preparation, and peace of mind for families and physicians. Simply put, PreSeek™ is the most comprehensive single gene cell-free fetal DNA screen available.




Thank‘s !