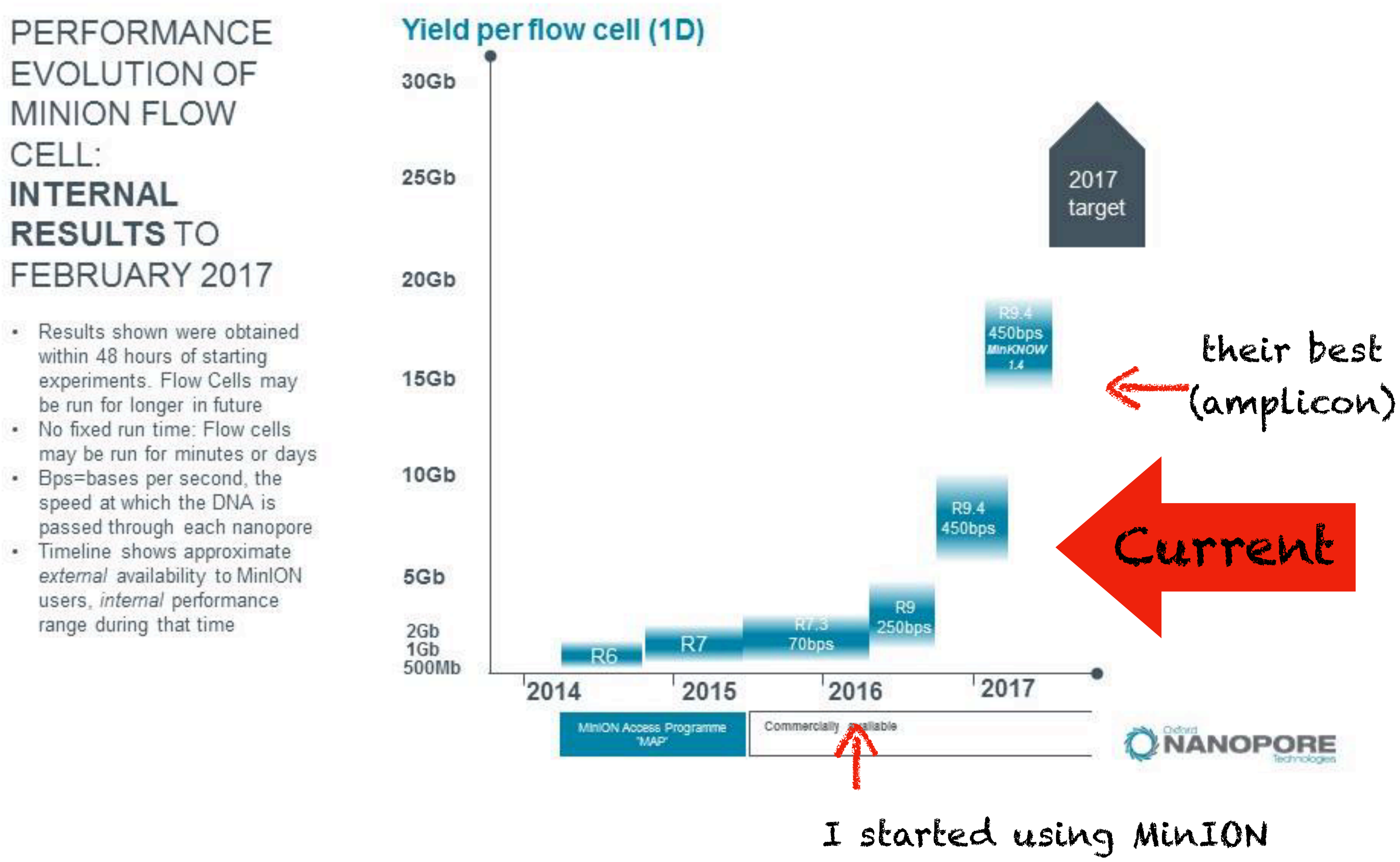
Medical Applications of Nanopore sequencing

Oxford Nanopore's MinION

Department of Human Genetics
Yokohama City University, Yokohama, Japan

Satomi Mitsuhashi, M.D., Ph.D.

Performance is improving



PERFORMANCE EVOLUTION OF MINION FLOW CELL: **INTERNAL RESULTS** TO FEBRUARY 2017

- Results shown were obtained within 48 hours of starting experiments. Flow Cells may be run for longer in future
- No fixed run time: Flow cells may be run for minutes or days
- Bps=bases per second, the speed at which the DNA is passed through each nanopore
- Timeline shows approximate *external* availability to MinION users, *internal* performance range during that time

their best (amplicon)

Current

I started using MinION

Today's topic

→ Examples of medical applications

→ Rapid diagnosis of infectious disease

→ D4Z4 Repeat sequencing
for the diagnosis
of a muscular dystrophy

Why MinION ?

Three major reasons to use MinION in medical (diagnostic) settings.

1. Portable

2. Rapid

3. Long read

1. Portable MinION enables “on-site” sequencing

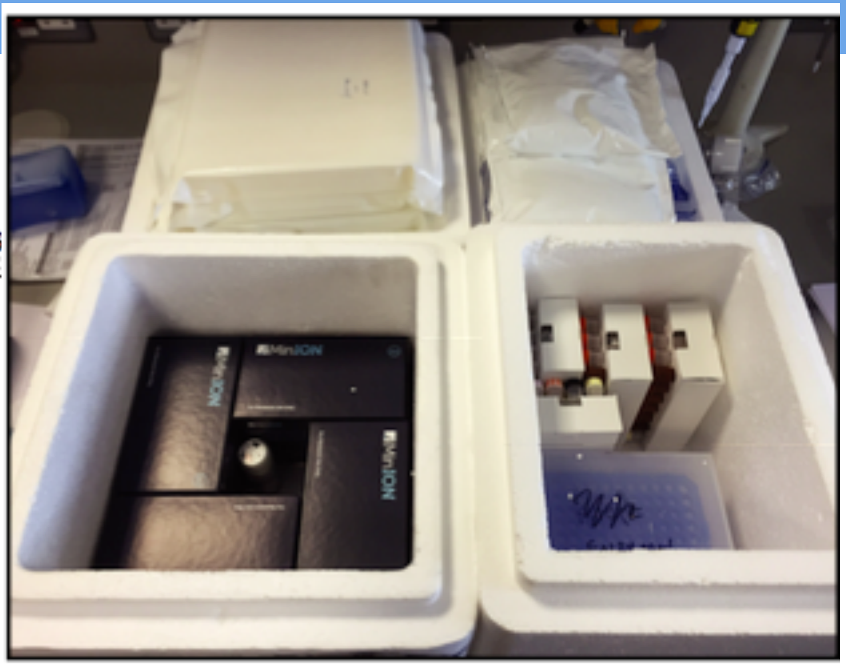
1. Portable

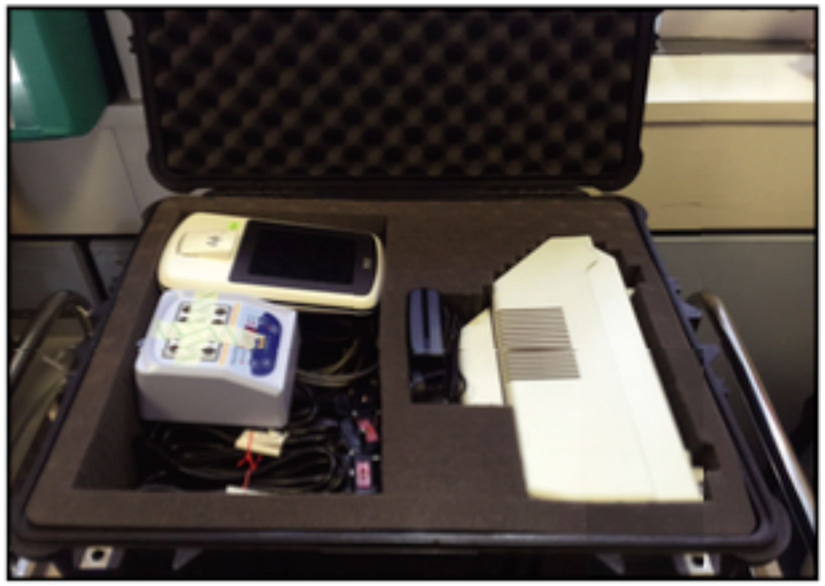
LETTER

doi:10.1038/nature16

Real-time, portable genome sequencing for Ebola surveillance

Joshua Quick[1]*, Nicholas J. Loman[1]*, Sophie Duraffour[2,3]*, Jared T. Simpson[4,5]*, Ettore Severi[6]*, Lauren Cowley[7]*,
Joseph Akoi Bore[2], Raymond Koundouno[2], Gytis Dudas[8], Amy Mikhail[7], Nobila Ouédraogo[9], Babak Afrough[2,10],
Amadou Bah[2,11], Jonathan H. J. Baum[2,3], Beate Becker-Ziaja[2,3], Jan Peter Boett
Álvaro Camino-Sánchez[2], Lisa L. Carter[2,13], Juliane Doerrbecker[2,3], Theresa E
Nicole Hetzelt[2,12], Julia Hinzmann[2,12], Tobias Holm[2,3], Liana Eleni Kafetzopoul
Eeva Kuisma[2,10], Christopher H. Logue[2,10], Antonio Mazzarelli[2,19], Sarah Meise
Didier Ngabo[2,10], Katja Nitzsche[2,3], Elisa Pallasch[2,3], Livia Victoria Patrono[2,3],
Natasha Y. Rickett[2,15,23], Andreas Sachse[2,12], Katrin Singethan[2,24], Inês Vitoria
Elsa G. Zekeng[2,15,23], Trina Racine[25], Alexander Bello[25], Amadou Alpha Sall[26],
N'Faly Magassouba[27], Cecelia V. Williams[28,29], Victoria Amburgey[28,29], Linda Wi
Frank Washington[29,30], Vanessa Monteil[31], Marine Jourdain[31], Marion Bererd[31],
Abdoulaye Camara[31], Marianne Gerard[31], Guillaume Bado[31], Bernard Baillet[31], D
Abdoulaye Diarra[34], Yacouba Savane[34], Raymond Bernard Pallawo[34], Giovanna
Isabelle Roger[34], Christopher J. Williams[6,37], Facinet Yattara[17], Kuiama Lewand
Daniel J. Turner[39], Georgios Pollakis[15,23], Julian A. Hiscox[15,23], David A. Matthew
Andrew McD. Johnston[41], Duncan Wilson[41], Emma Hutley[42], Erasmus Smit[43],
Kilian Stoecker[2,44], Erna Fleischmann[2,44], Martin Gabriel[2,3], Simon A. Weller[38],
Sakoba Keïta[17], Andrew Rambaut[8,46,47], Pierre Formenty[34], Stephan Günther[2,3]



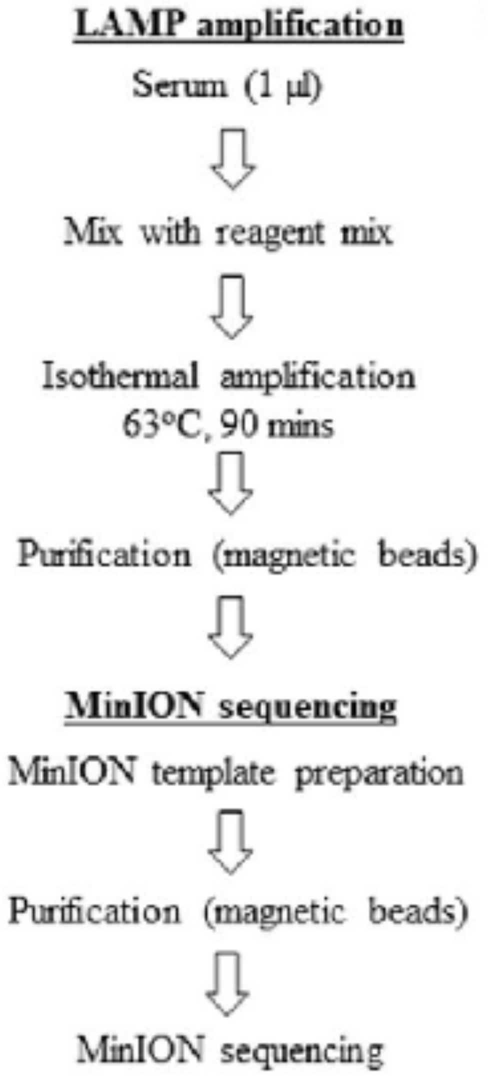


2. Rapid

Dengue serotyping within one day

Serotyping dengue virus with isothermal amplification and a portable sequencer

Junya Yamagishi[1,2], Lucky R. Runtuwene[3], Kyoko Hayashida[1,4], Arthur E. Mongan[5], Lan Anh Nguyen Thi[6], Linh Nguyen Thuy[6], Cam Nguyen Nhat[7], Kriengsak Limkittikul[8], Chukiat Sirivichayakul[8], Nuankanya Sathirapongsasuti[9], Martin Frith[10], Wojciech Makalowski[11], Yuki Eshita[12], Sumio Sugano[3] & Yutaka Suzuki[3]



LAMP amplification

Serum (1 µl)

⇩

Mix with reagent mix

⇩

Isothermal amplification
63°C, 90 mins

⇩

Purification (magnetic beads)

⇩

MinION sequencing

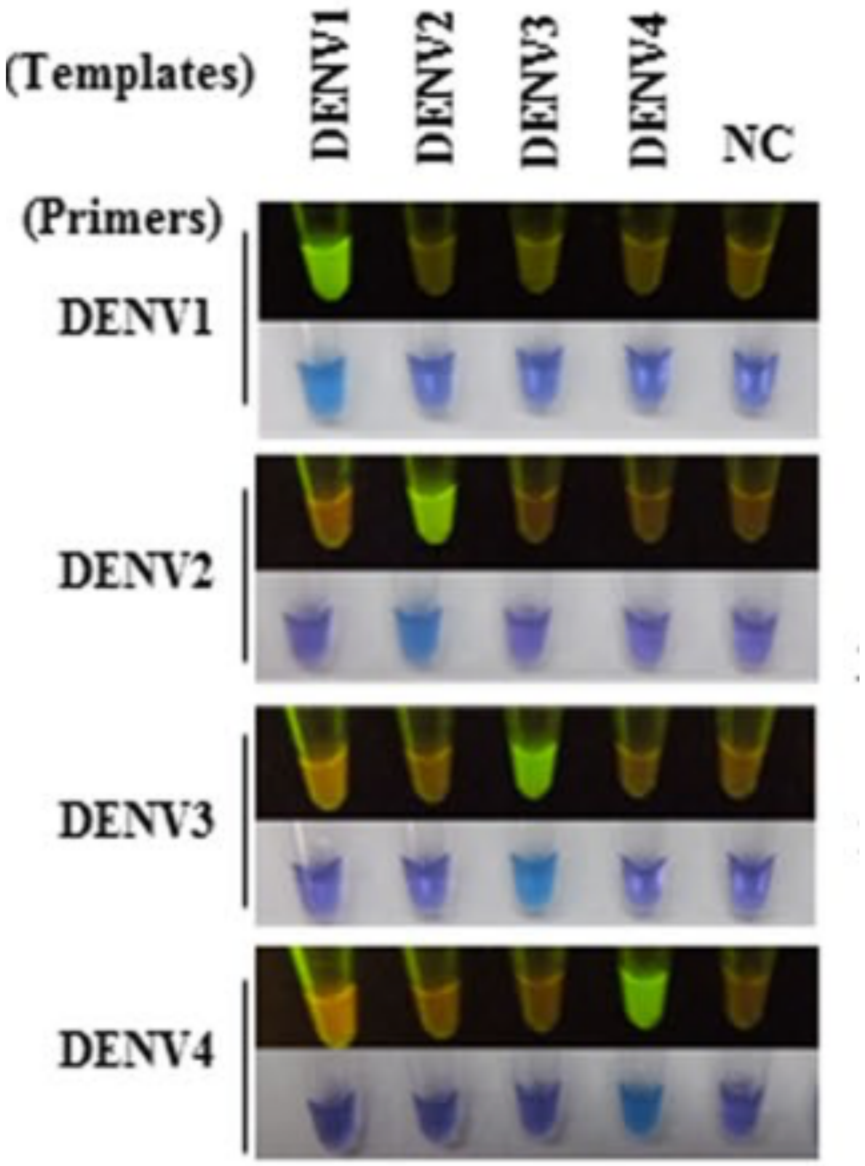
MinION template preparation

⇩

Purification (magnetic beads)

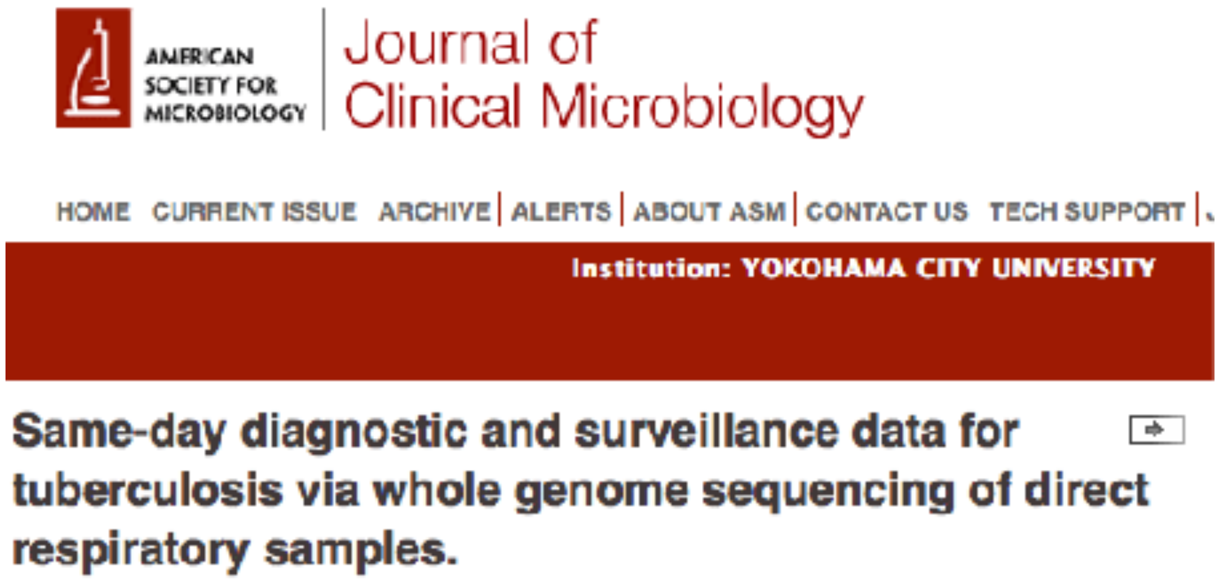
⇩

MinION sequencing



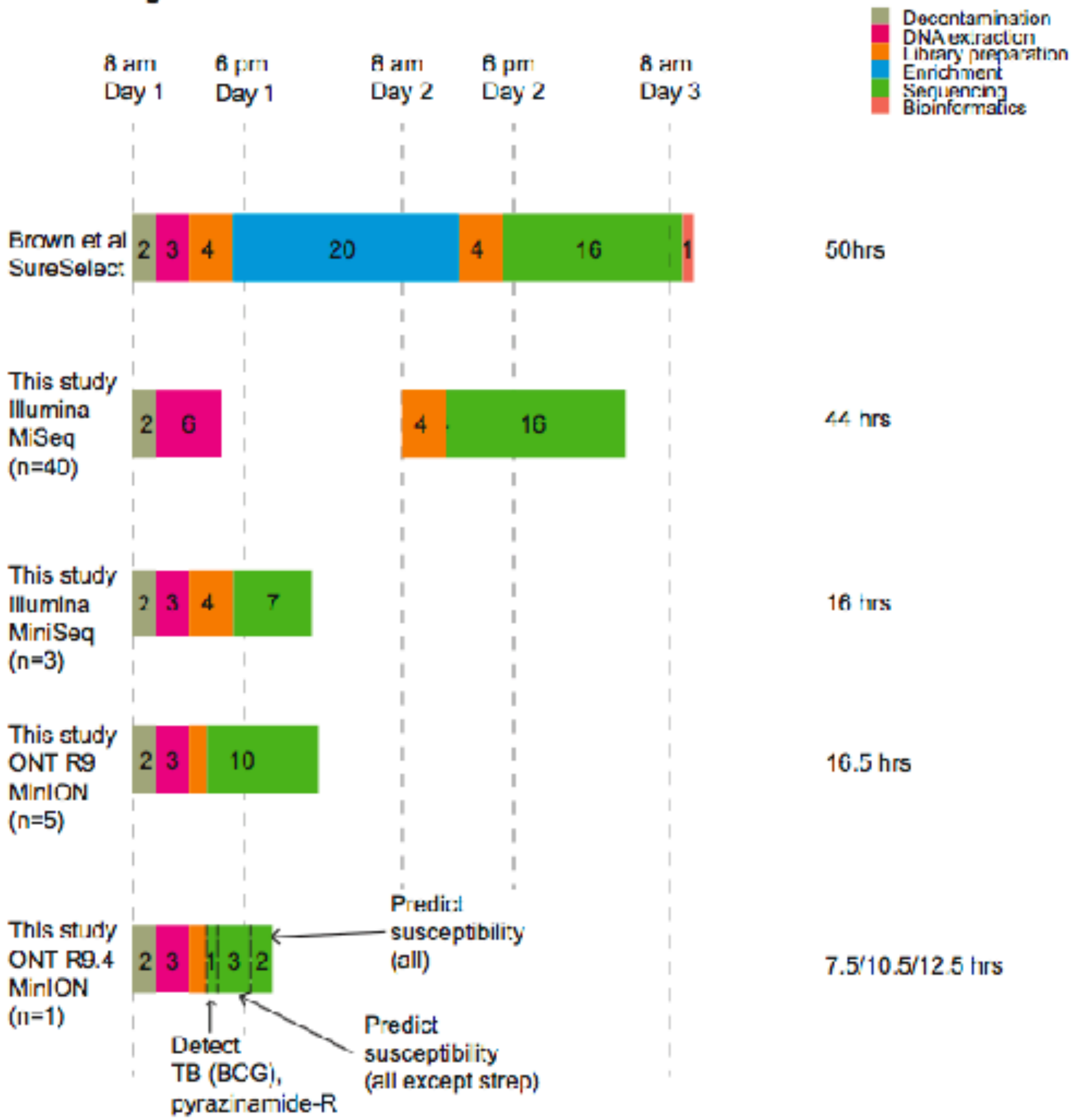
(Yamagishi et al. 2017. Sci Rep)

2. Rapid Tuberculosis diagnosis within one day



Same-day diagnostic and surveillance data for tuberculosis via whole genome sequencing of direct respiratory samples.

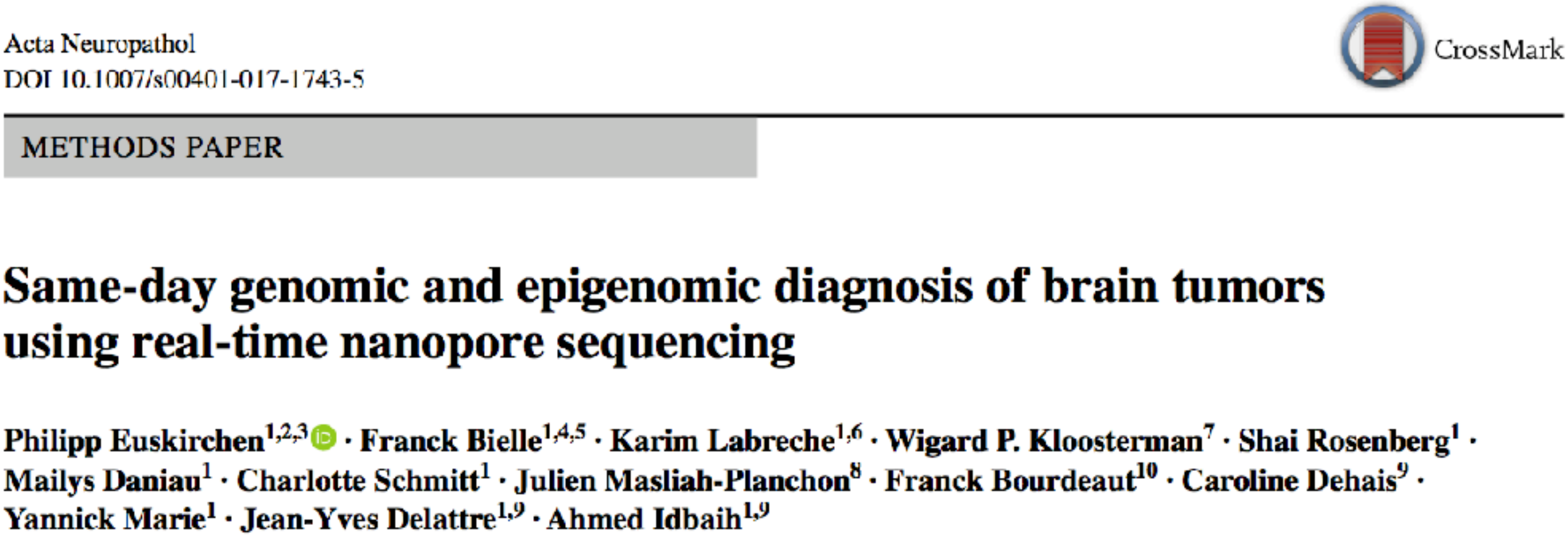
Figure 5



(Votintseva et al. 2017. J Clin Microbiol)

2. Rapid

Brain tumor diagnosis within one day

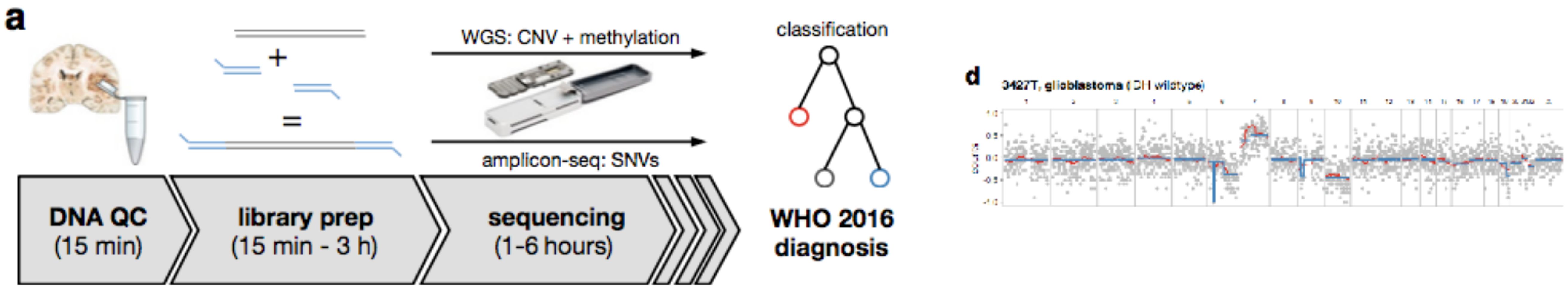


Acta Neuropathol
DOI 10.1007/s00401-017-1743-5

METHODS PAPER

Same-day genomic and epigenomic diagnosis of brain tumors using real-time nanopore sequencing

Philipp Euskirchen[1,2,3] · Franck Bielle[1,4,5] · Karim Labreche[1,6] · Wigard P. Kloosterman[7] · Shai Rosenberg[1] · Mailys Daniau[1] · Charlotte Schmitt[1] · Julien Masliah-Planchon[8] · Franck Bourdeaut[10] · Caroline Dehais[9] · Yannick Marie[1] · Jean-Yves Delattre[1,9] · Ahmed Idbaih[1,9]



(Euskirchen et al. 2017. Acta Neuropathol)

2. Rapid sequencing kit

Rapid sequencing kit (SQK-RAD002)

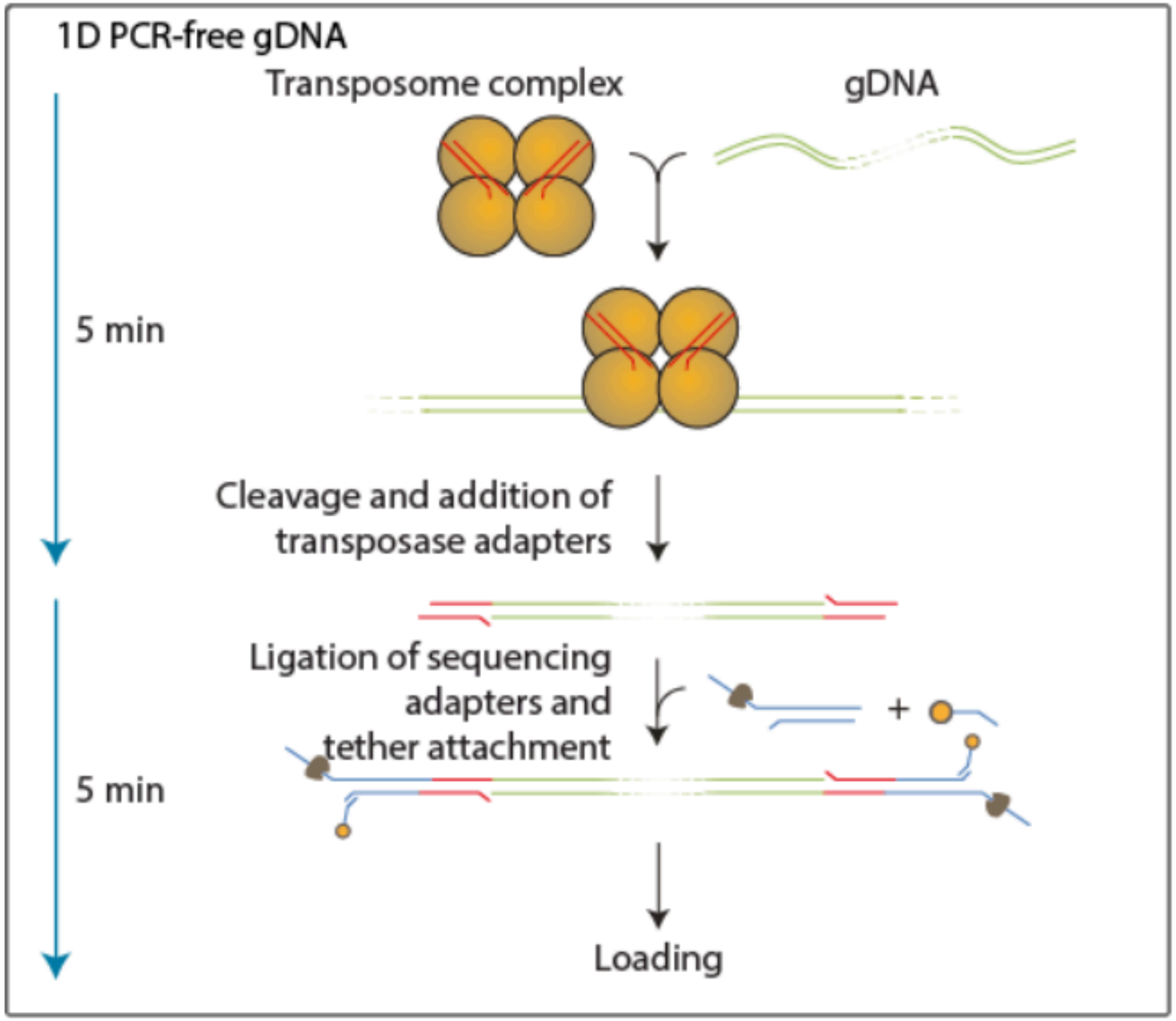


Figure RSWF: Workflow of the Rapid Sequencing library preparation protocol.

< 10min

Add 2.5ul FRM to 200ng of DNA
incubate for 1min at 30°C
incubate for 1min at 75°C

Add 1ul RAD / 0.2ul TA ligase
incubate for 5min at RT

↓

Ready for sequencing!

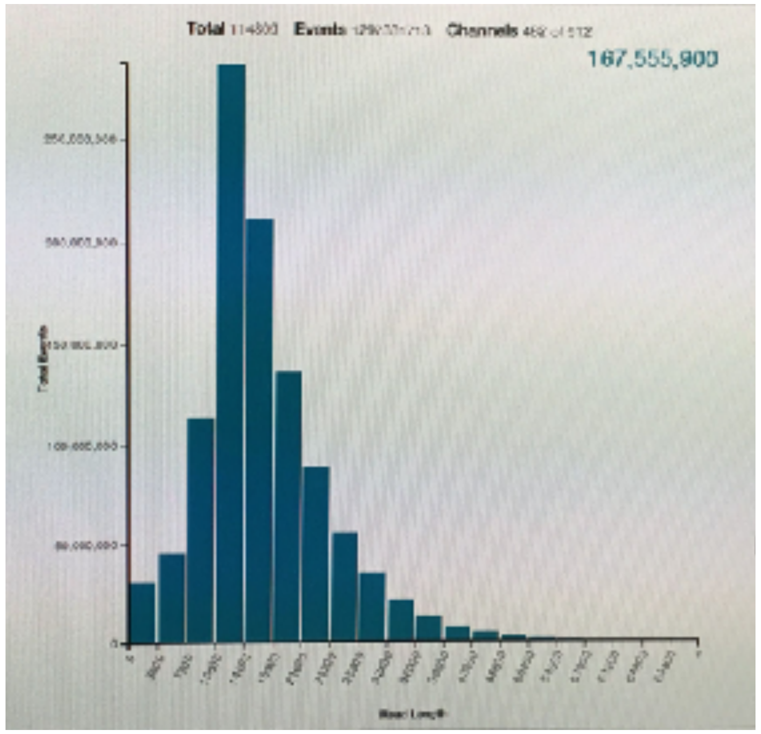
3. Long read

short read vs. long read

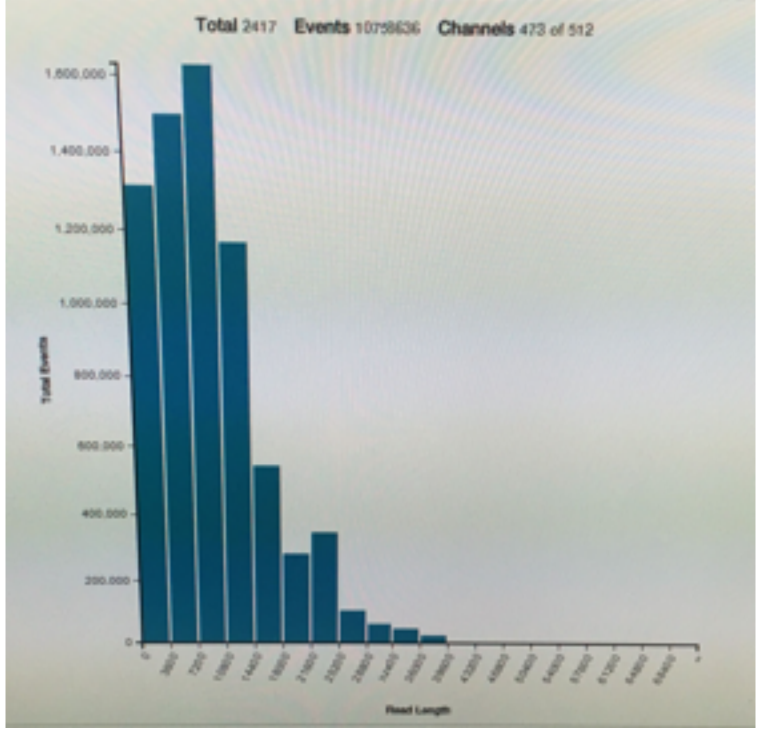
↓

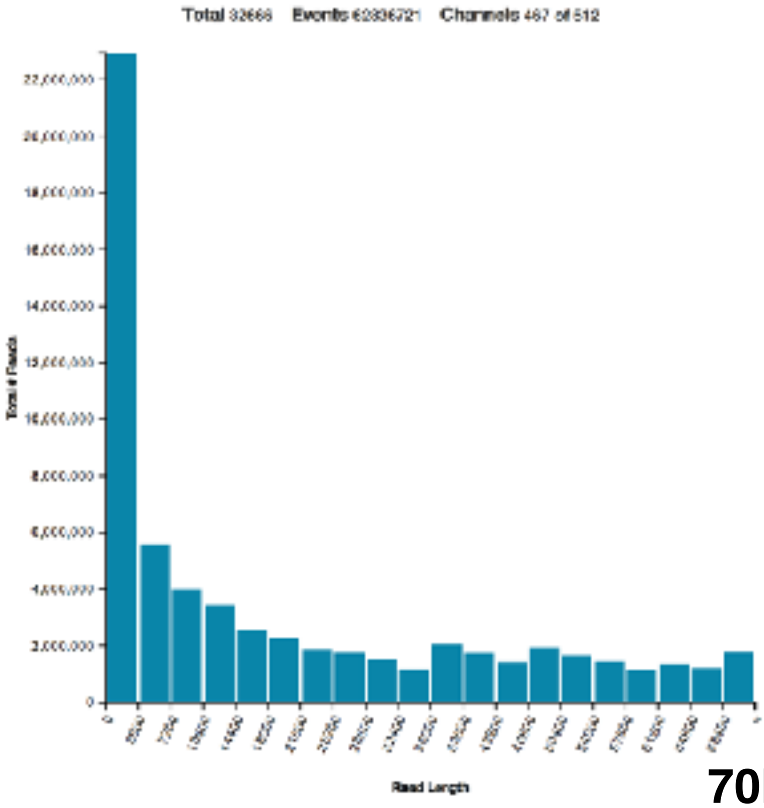
sprint vs. full marathon

14k



~20k





70k

3. Long read

Difficult / impossible regions to sequence in human genome using short read sequencers

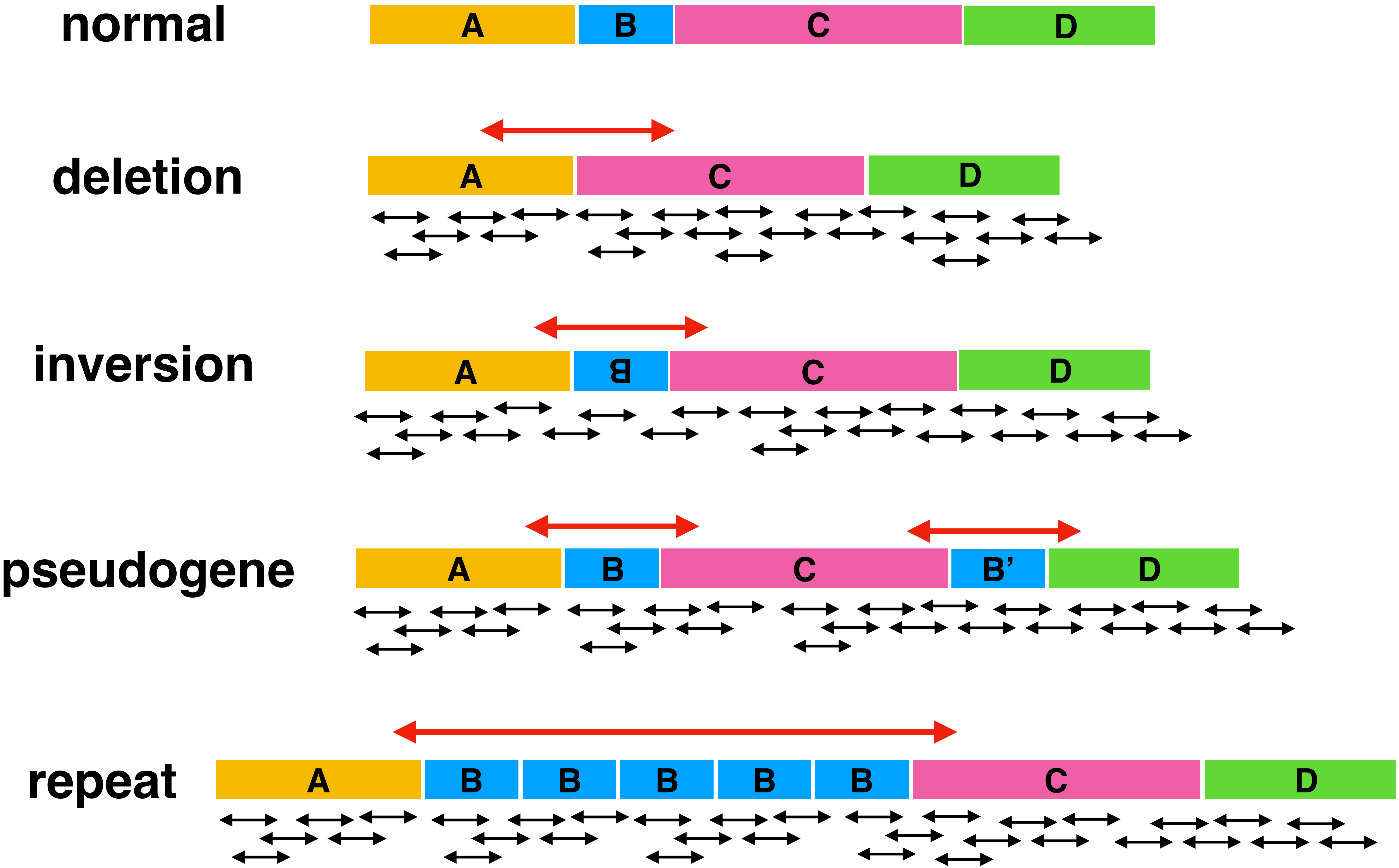
Genomic structural rearrangements

Segmental duplications

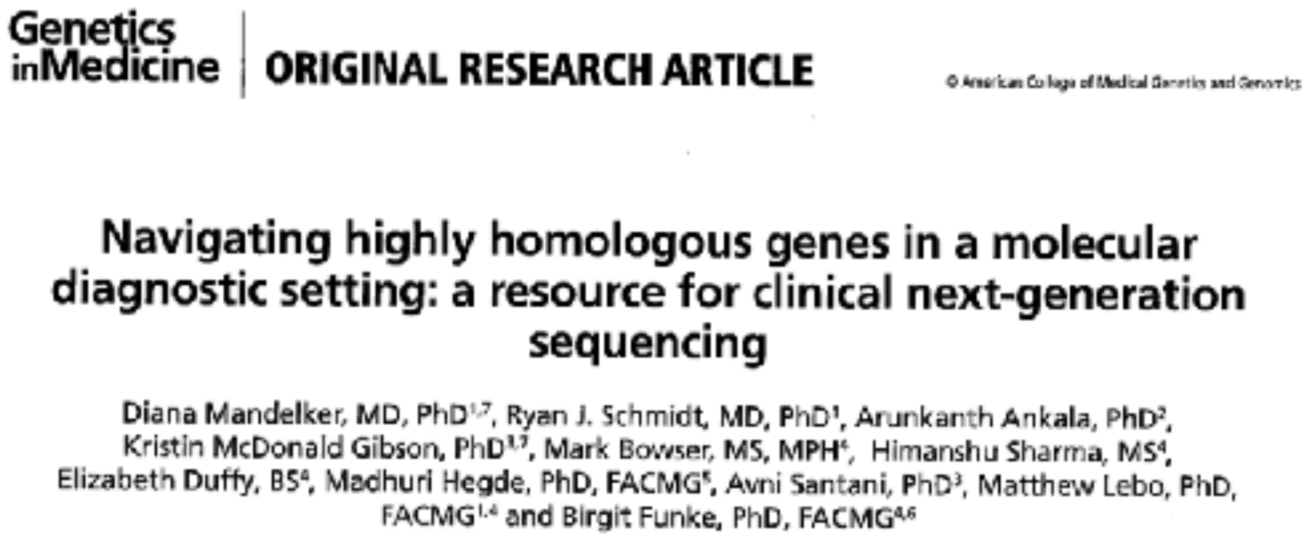
Macrosatellite repeats

etc…

long read *vs.* short read



Short read sequence does not cover whole genome

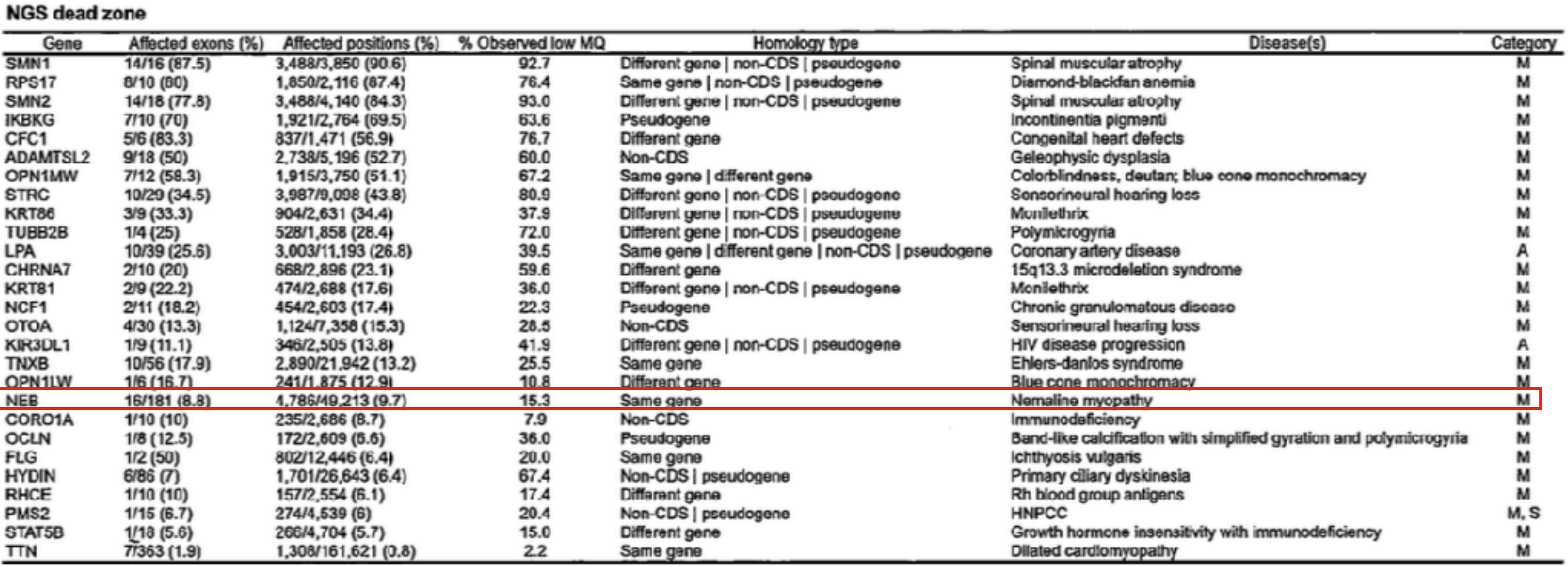


Genetics inMedicine | ORIGINAL RESEARCH ARTICLE

Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing

Diana Mandelker, MD, PhD[1,7], Ryan J. Schmidt, MD, PhD[1], Arunkanth Ankala, PhD[2],
Kristin McDonald Gibson, PhD[3,7], Mark Bowser, MS, MPH[4], Himanshu Sharma, MS[4],
Elizabeth Duffy, BS[4], Madhuri Hegde, PhD, FACMG[5], Avni Santani, PhD[3], Matthew Lebo, PhD,
FACMG[1,4] and Birgit Funke, PhD, FACMG[4,6]

“NGS dead” zone



NGS dead zone

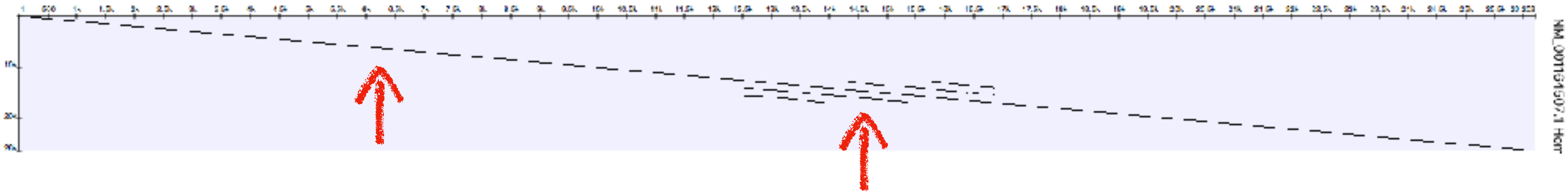
Gene	Affected exons (%)	Affected positions (%)	% Observed low MQ	Homology type	Disease(s)	Category
SMN1	14/16 (87.5)	3,488/3,850 (90.6)	92.7	Different gene | non-CDS | pseudogene	Spinal muscular atrophy	M
RPS17	8/10 (80)	1,850/2,116 (87.4)	76.4	Same gene | non-CDS | pseudogene	Diamond-blackfan anemia	M
SMN2	14/18 (77.8)	3,488/4,140 (84.3)	93.0	Different gene | non-CDS | pseudogene	Spinal muscular atrophy	M
IKBKG	7/10 (70)	1,921/2,764 (69.5)	63.6	Pseudogene	Incontinentia pigmenti	M
CFC1	5/6 (83.3)	837/1,471 (56.9)	76.7	Different gene	Congenital heart defects	M
ADAMTSL2	9/18 (50)	2,738/5,196 (52.7)	60.0	Non-CDS	Geleophysic dysplasia	M
OPN1MW	7/12 (58.3)	1,915/3,750 (51.1)	67.2	Same gene | different gene	Colorblindness, deutan; blue cone monochromacy	M
STRC	10/29 (34.5)	3,987/9,098 (43.8)	80.9	Different gene | non-CDS | pseudogene	Sensorineural hearing loss	M
KRT86	3/9 (33.3)	904/2,631 (34.4)	37.9	Different gene | non-CDS | pseudogene	Monilethrix	M
TUBB2B	1/4 (25)	528/1,858 (28.4)	72.0	Different gene | non-CDS | pseudogene	Polymicrogyria	M
LPA	10/39 (25.6)	3,003/11,193 (26.8)	39.5	Same gene | different gene | non-CDS | pseudogene	Coronary artery disease	A
CHRNA7	2/10 (20)	668/2,896 (23.1)	59.6	Different gene	15q13.3 microdeletion syndrome	M
KRT81	2/9 (22.2)	474/2,688 (17.6)	36.0	Different gene | non-CDS | pseudogene	Monilethrix	M
NCF1	2/11 (18.2)	454/2,603 (17.4)	22.3	Pseudogene	Chronic granulomatous disease	M
OTOA	4/30 (13.3)	1,124/7,358 (15.3)	28.5	Non-CDS	Sensorineural hearing loss	M
KIR3DL1	1/9 (11.1)	346/2,505 (13.8)	41.9	Different gene | non-CDS | pseudogene	HIV disease progression	A
TNXB	10/56 (17.9)	2,890/21,942 (13.2)	25.5	Same gene	Ehlers-danlos syndrome	M
OPN1LW	1/6 (16.7)	241/1,875 (12.9)	10.8	Different gene	Blue cone monochromacy	M
NEB	16/181 (8.8)	4,786/49,213 (9.7)	15.3	Same gene	Nemaline myopathy	M
CORO1A	1/10 (10)	235/2,686 (8.7)	7.9	Non-CDS	Immunodeficiency	M
OCLN	1/8 (12.5)	172/2,609 (6.6)	36.0	Pseudogene	Band-like calcification with simplified gyration and polymicrogyria	M
FLG	1/2 (50)	802/12,446 (6.4)	20.0	Same gene	Ichthyosis vulgaris	M
HYDIN	6/86 (7)	1,701/26,643 (6.4)	67.4	Non-CDS | pseudogene	Primary ciliary dyskinesia	M
RHCE	1/10 (10)	157/2,554 (6.1)	17.4	Different gene	Rh blood group antigens	M
PMS2	1/15 (6.7)	274/4,539 (6)	20.4	Non-CDS | pseudogene	HNPCC	M, S
STAT5B	1/18 (5.6)	266/4,704 (5.7)	15.0	Different gene	Growth hormone insensitivity with immunodeficiency	M
TTN	7/363 (1.9)	1,308/161,621 (0.8)	2.2	Same gene	Dilated cardiomyopathy	M

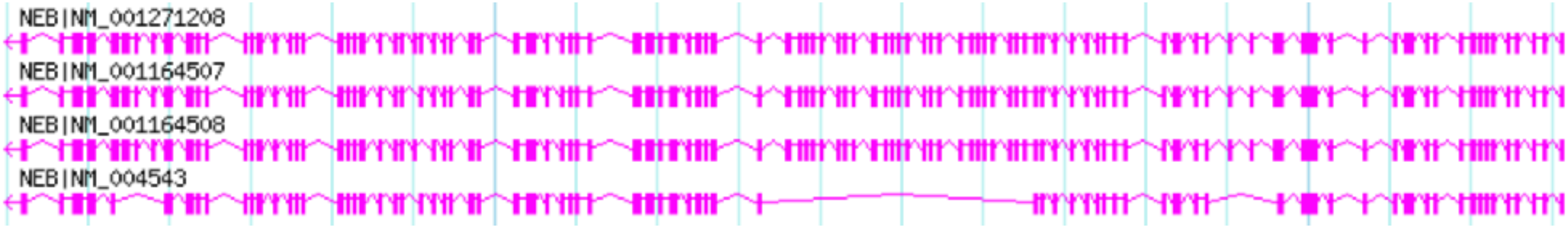
Sequencing regions of high sequence homology

(Mandelker et al. 2016 Genet Med)

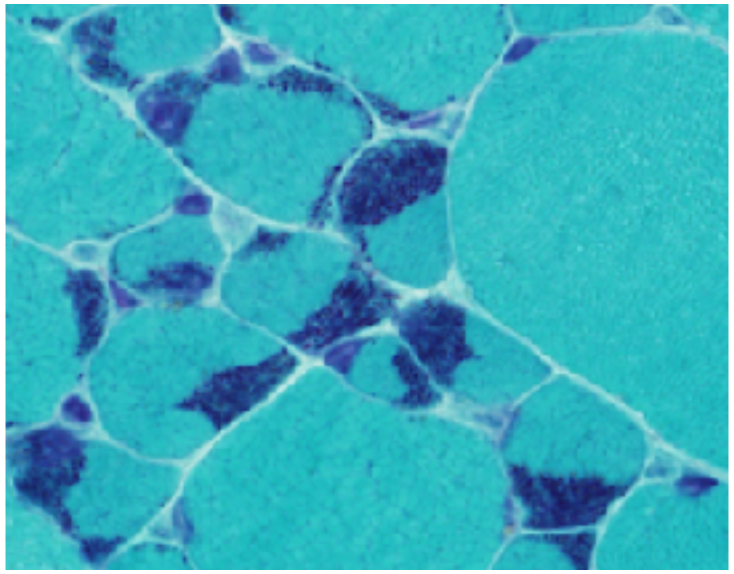
Example

NEB : nebuline

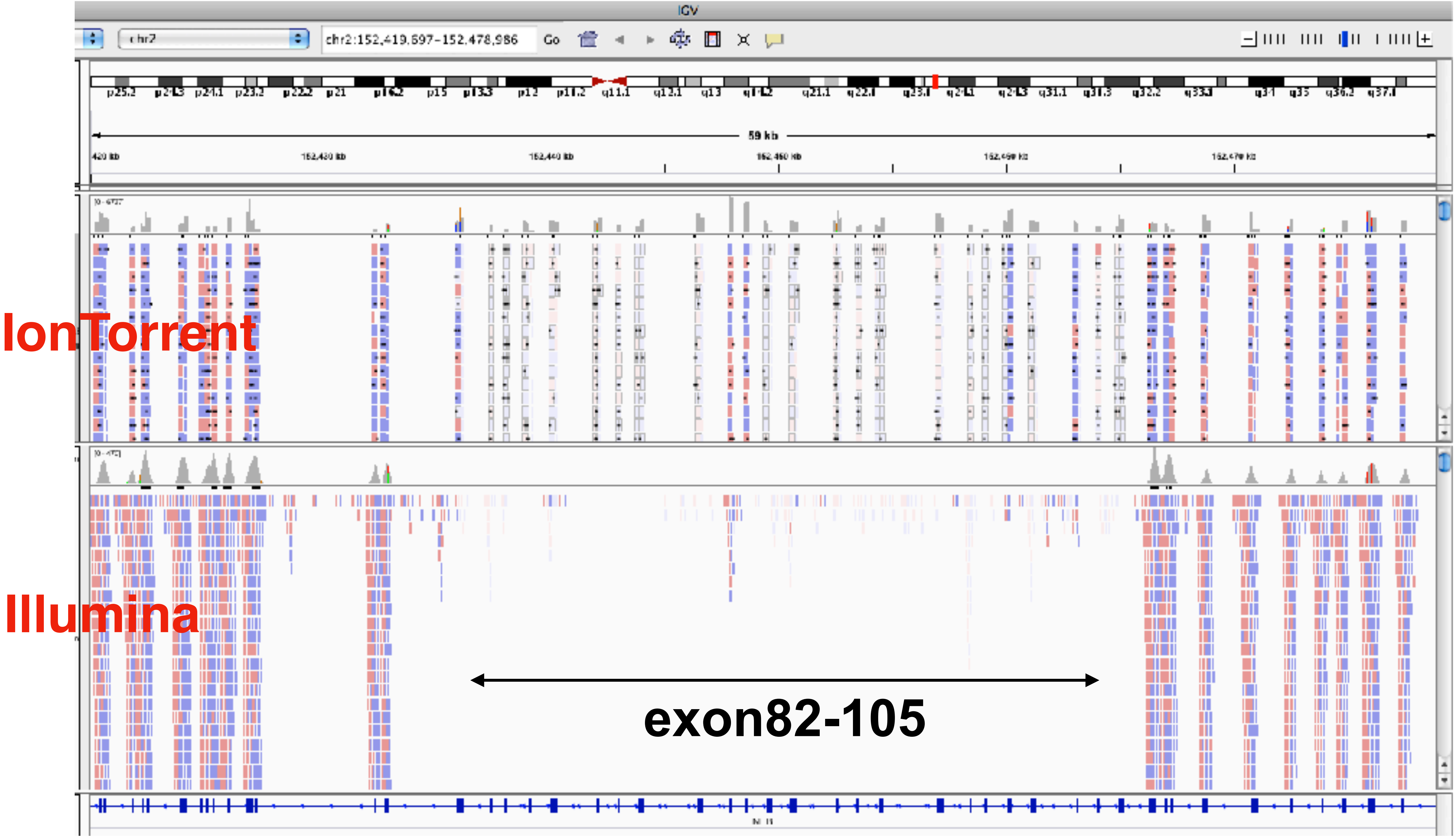




Recessive mutations in *NEB* cause congenital myopathy

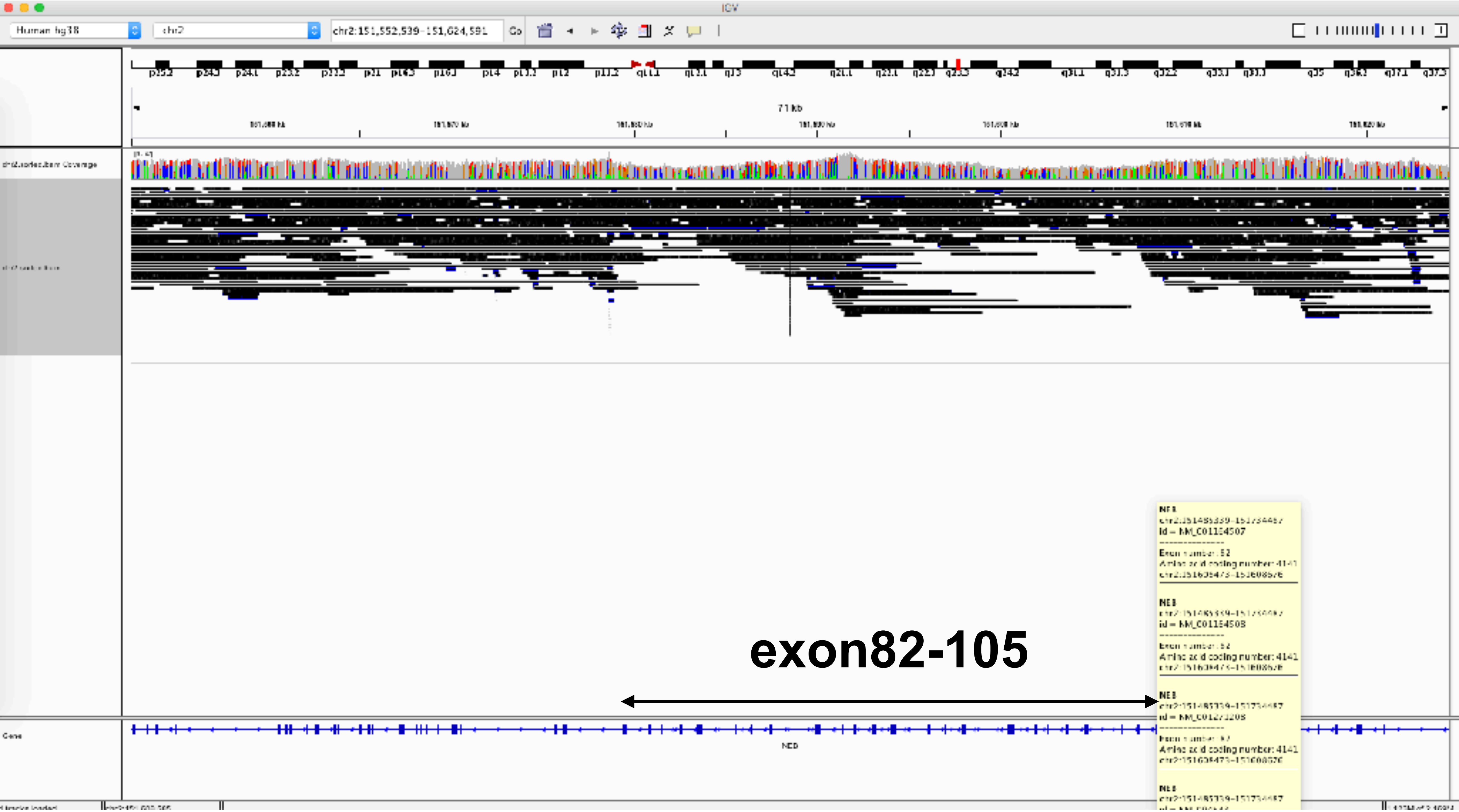


Target-sequencing using Illumina and IonTorrent



NEB

MinION sequencing



NEB
chr2:151485339-151734487
id = NM_001164507

Exon number: 82
Amino acid coding number: 4141
chr2:151608473-151608576

NEB
chr2:151485339-151734487
id = NM_001164508

Exon number: 82
Amino acid coding number: 4141
chr2:151608473-151608576

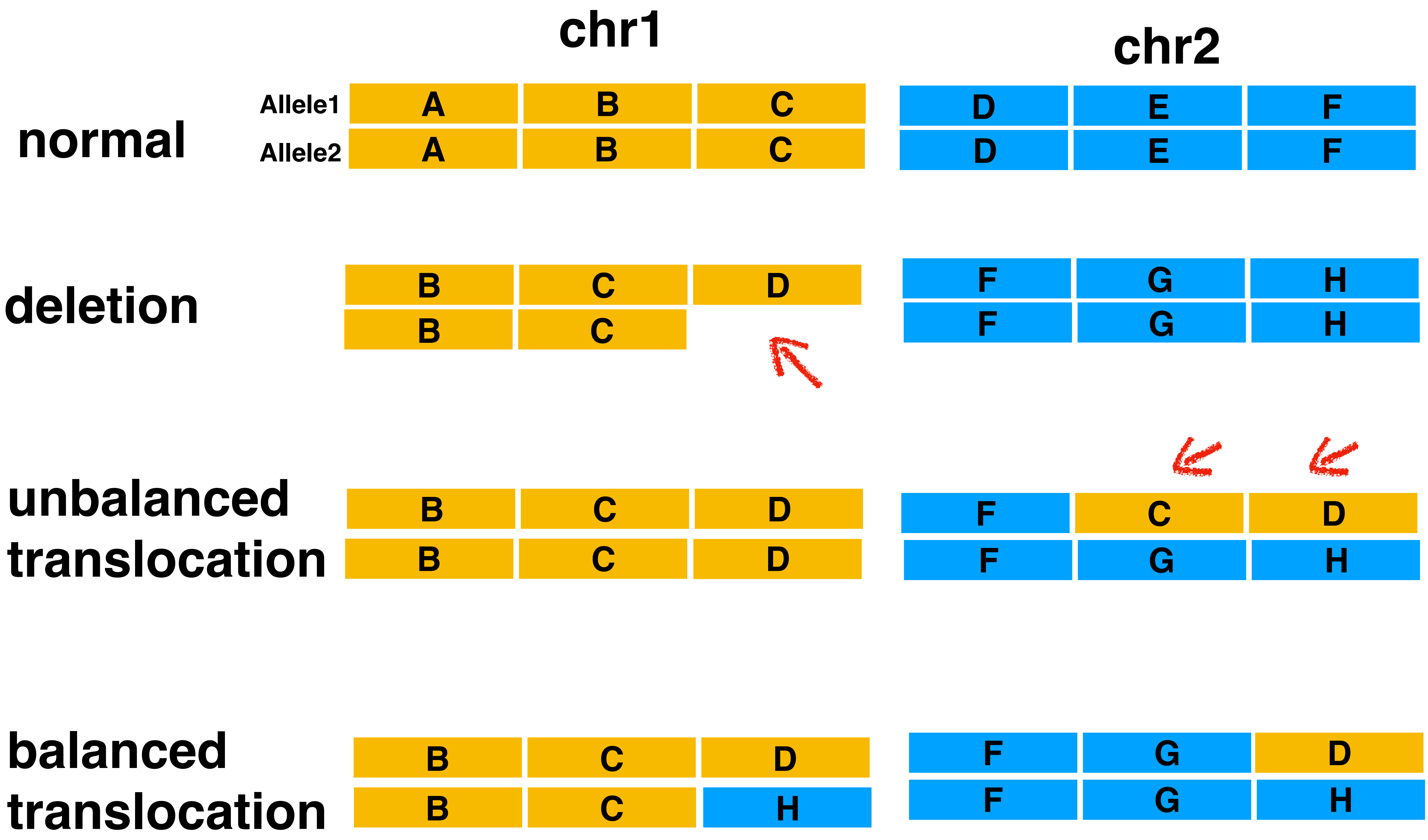
NEB
chr2:151485339-151734487
id = NM_001271208

Exon number: 82
Amino acid coding number: 4141
chr2:151608473-151608576

NEB
chr2:151485339-151734487

exon82-105

Chromosomal rearrangements

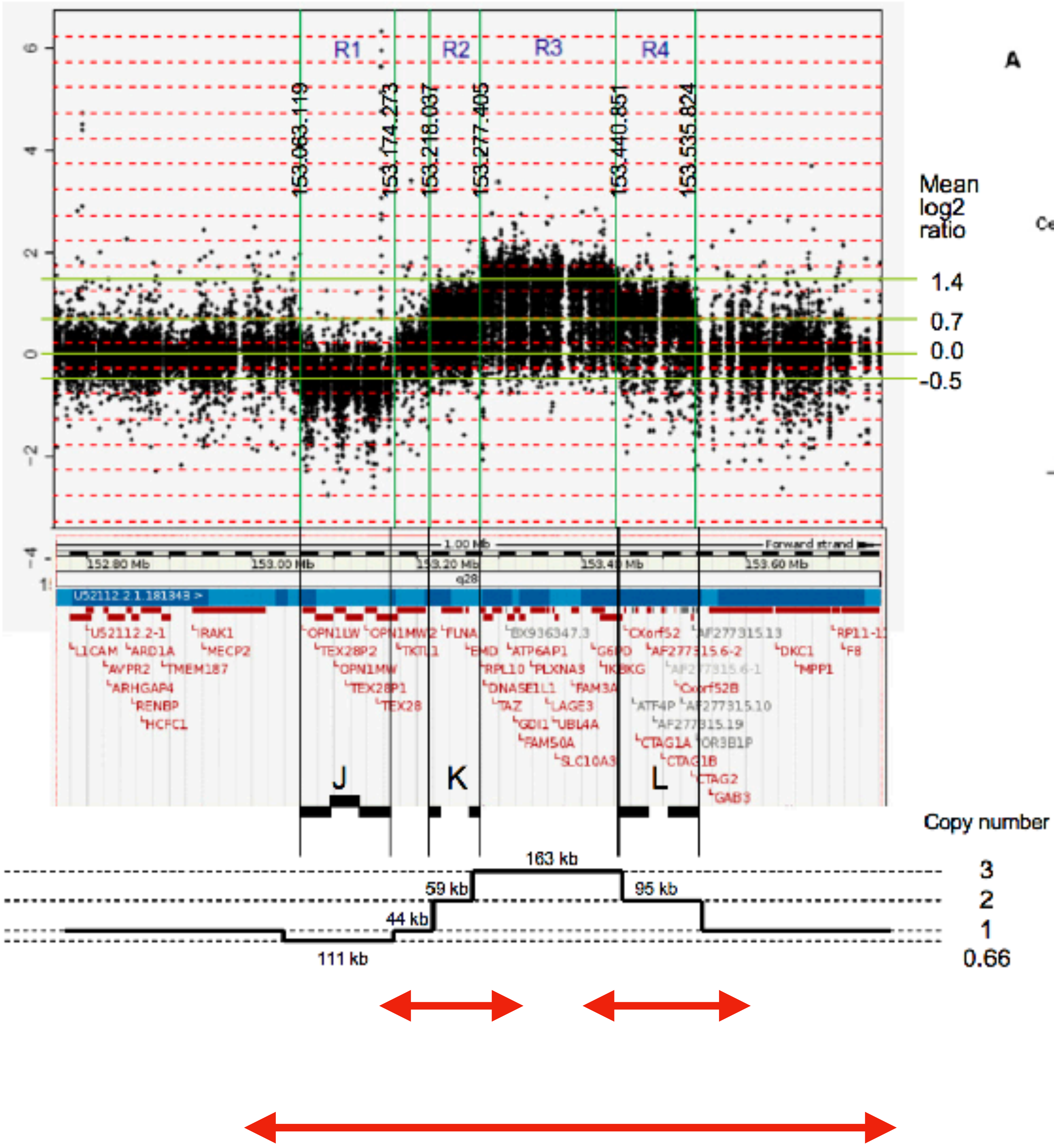


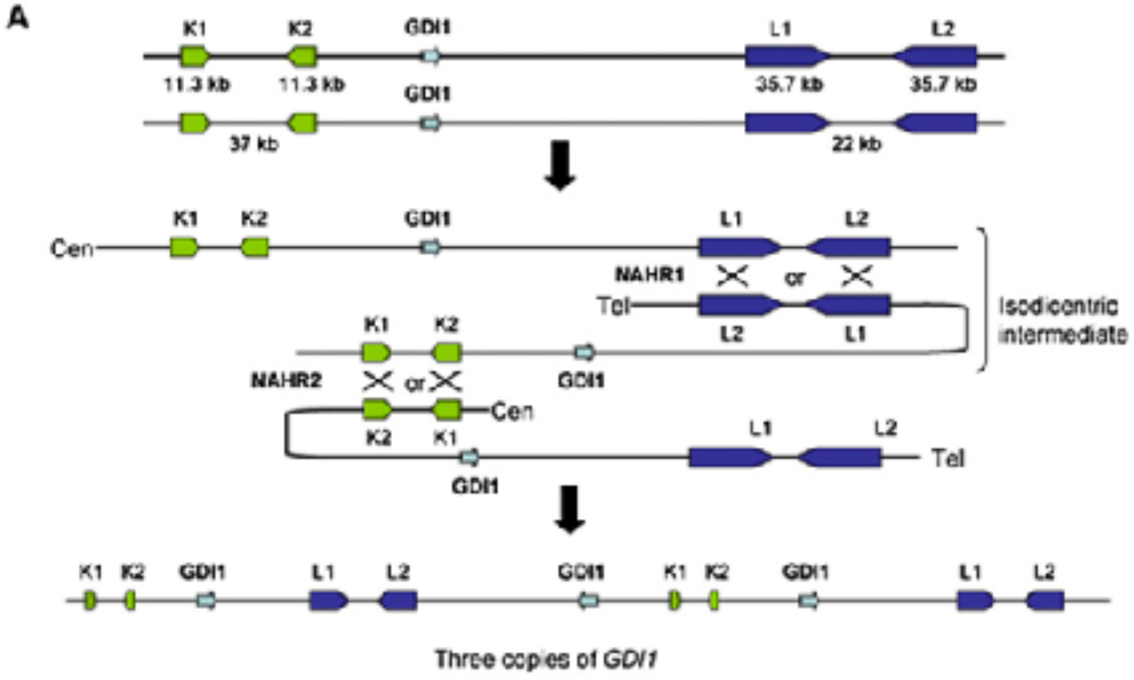
		chr1			chr2	
normal	Allele1	A	B	C	D	E	F
	Allele2	A	B	C	D	E	F
deletion		B	C	D	F	G	H
		B	C		F	G	H
unbalanced translocation		B	C	D	F	C	D
		B	C	D	F	G	H
balanced translocation		B	C	D	F	G	D
		B	C	H	F	G	H

Chromosomal rearrangements

chrX

Array CGH

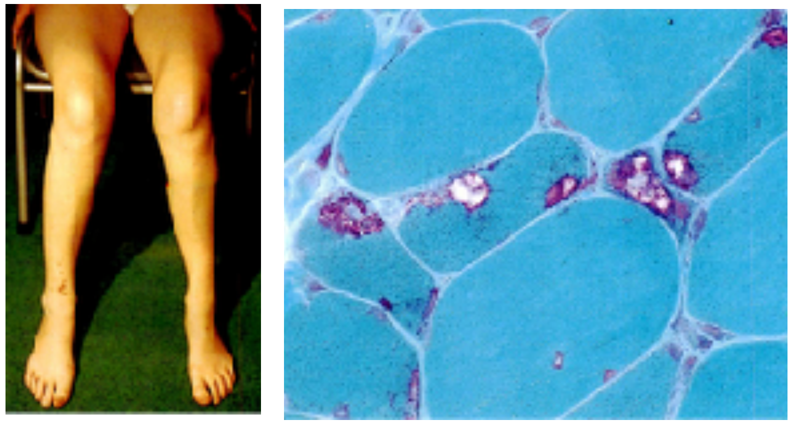




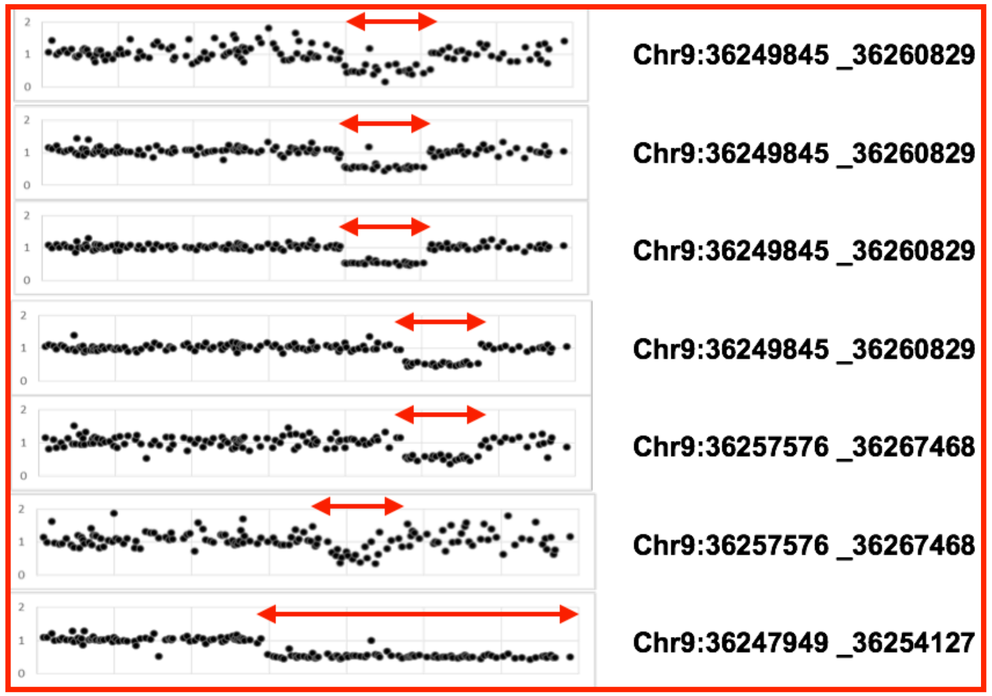
(Vandewalle et al 2009 Am J Hum Genet)

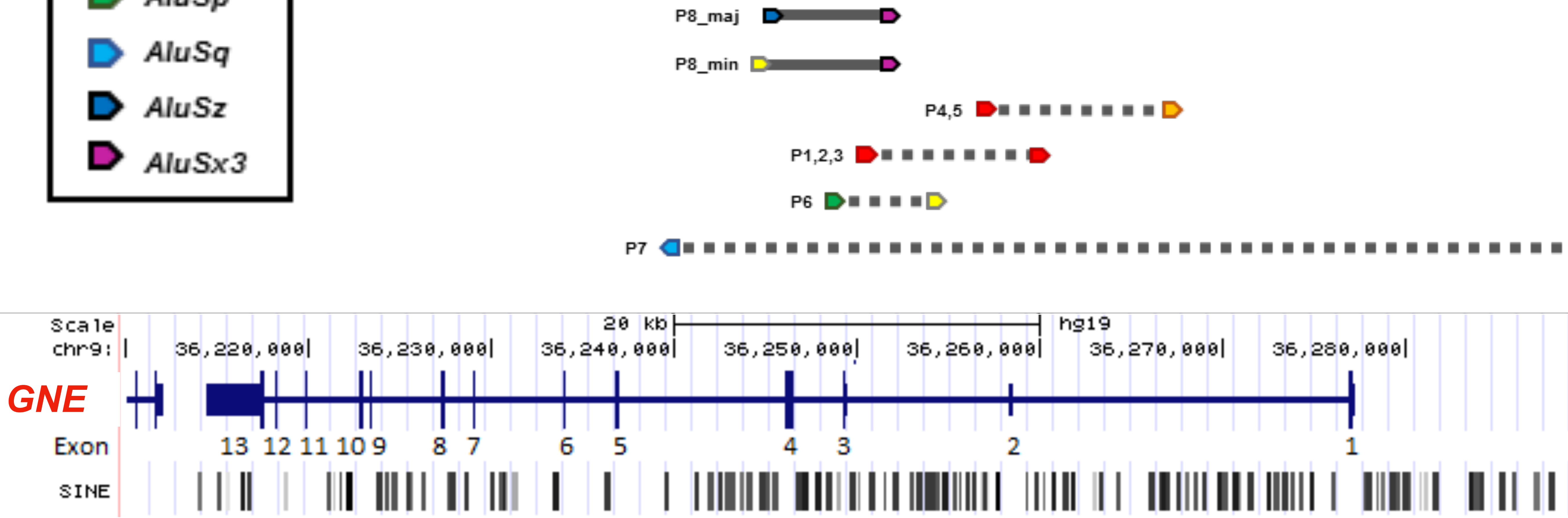
Breakpoints are prone to be at repetitive sequences

***GNE* myopathy**



(Nonaka et al. 1980 J Neurol Sci.)

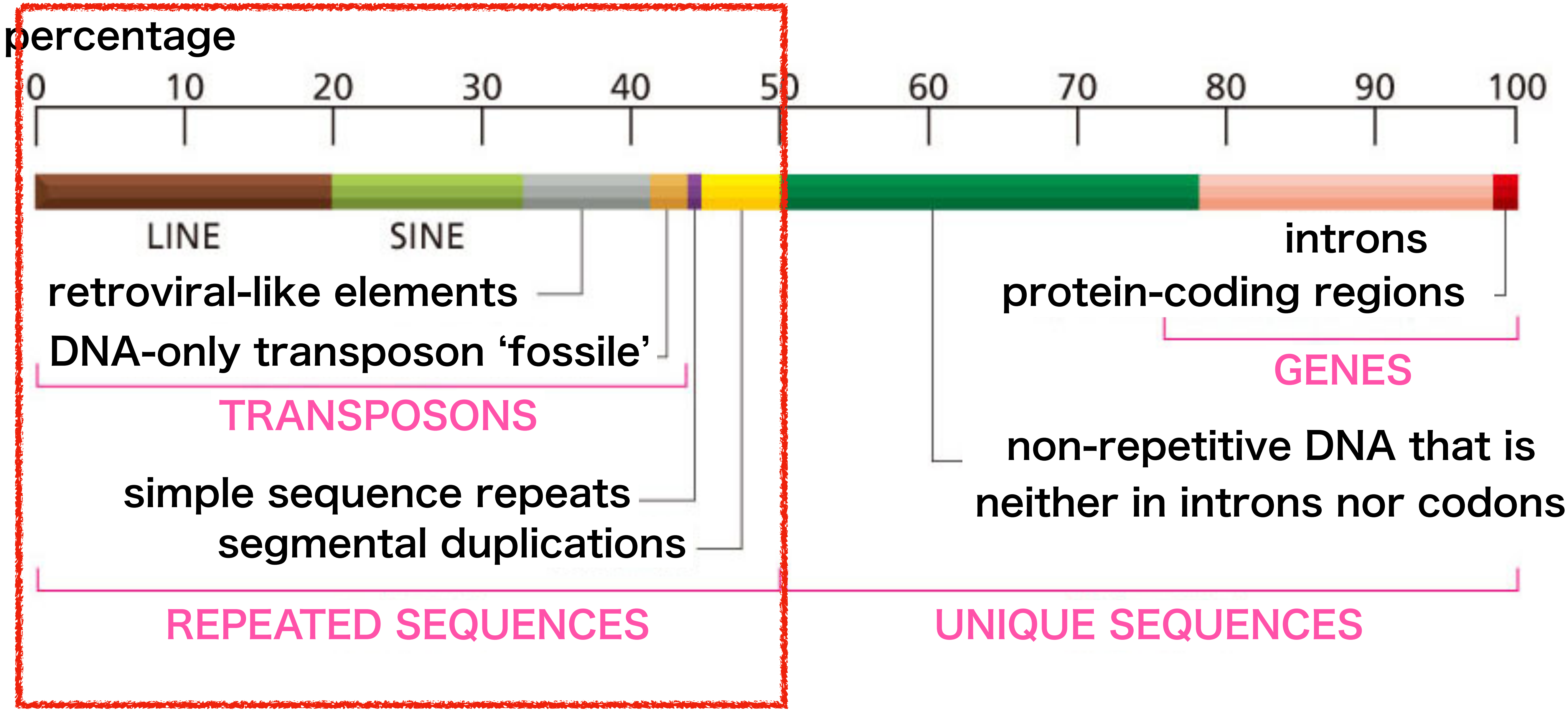




(Zhu et al. 2016 J Hum Genet)

Overview of Human Genome Contents

Molecular Biology of THE CELL 5th Edition



retroviral-like elements

DNA-only transposon ‘fossile’

TRANSPOSONS

simple sequence repeats

segmental duplications

REPEATED SEQUENCES

introns

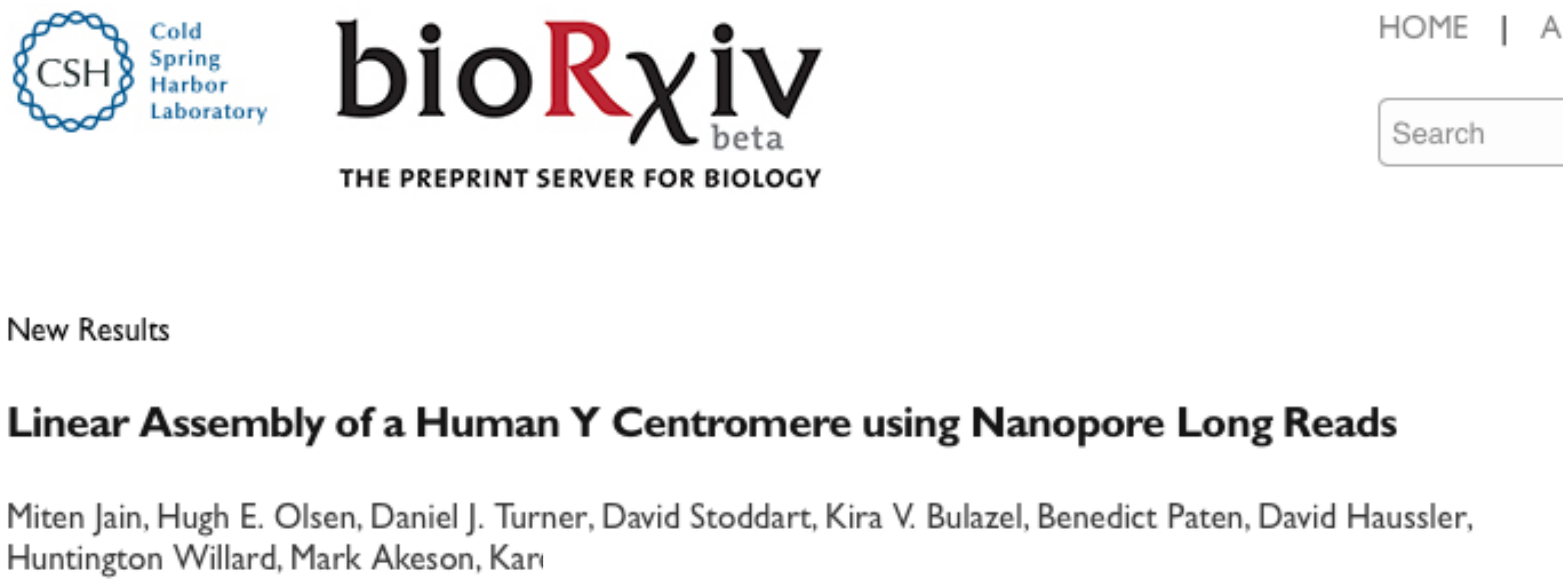
protein-coding regions

GENES

non-repetitive DNA that is neither in introns nor codons

UNIQUE SEQUENCES

Centromere sequence

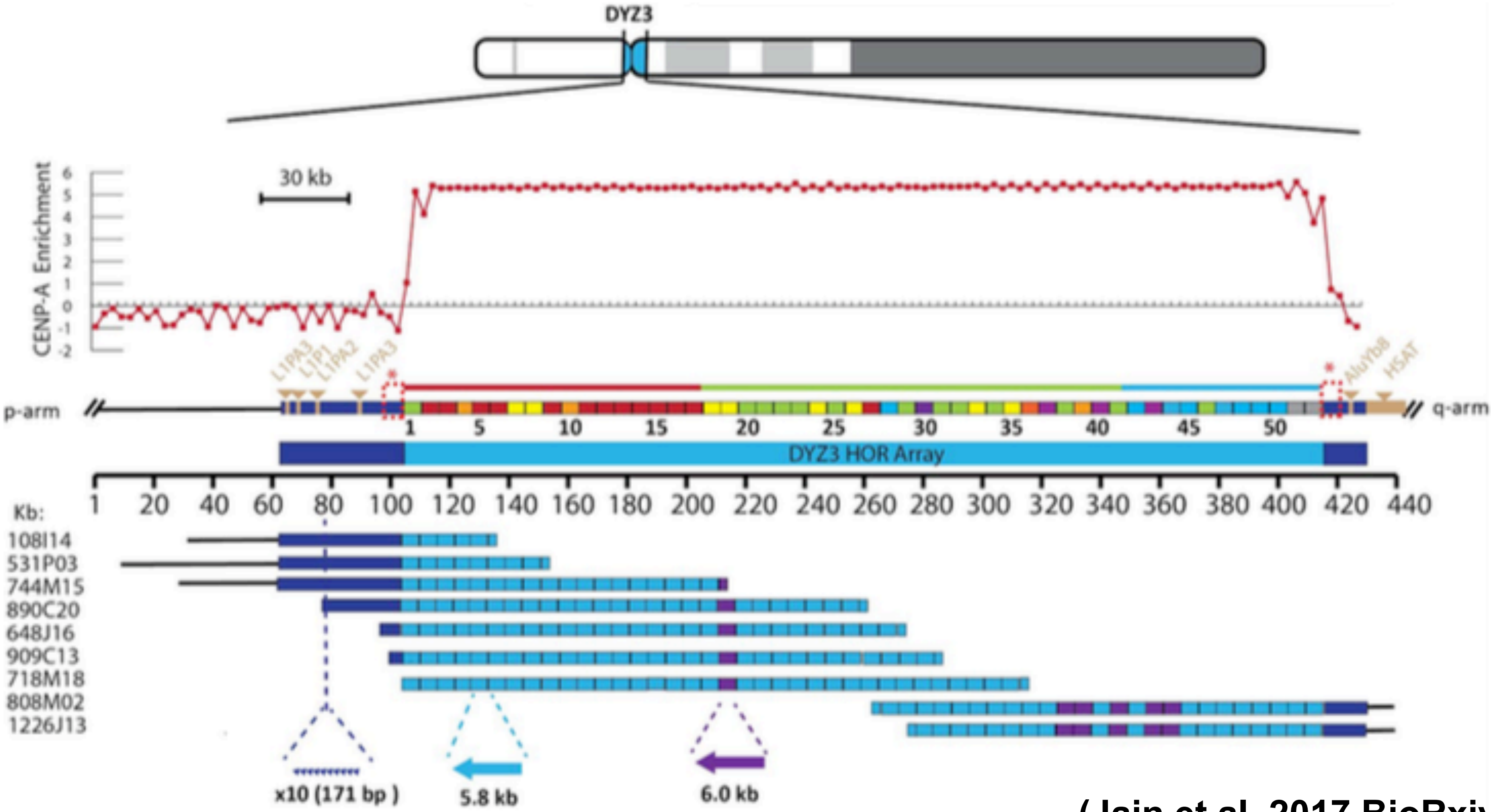


New Results

Linear Assembly of a Human Y Centromere using Nanopore Long Reads

Miten Jain, Hugh E. Olsen, Daniel J. Turner, David Stoddart, Kira V. Bulazel, Benedict Paten, David Haussler, Huntington Willard, Mark Akeson, Kar

doi: https://doi.org/10.1101/170373



(Jain et al. 2017 BioRxiv)

→ Rapid diagnosis of infectious disease

Rapid sequencing

(+ Portability, Long read)

Bacterial infection disease in local hospital

Aim: rapid and accurate identification of pathogen

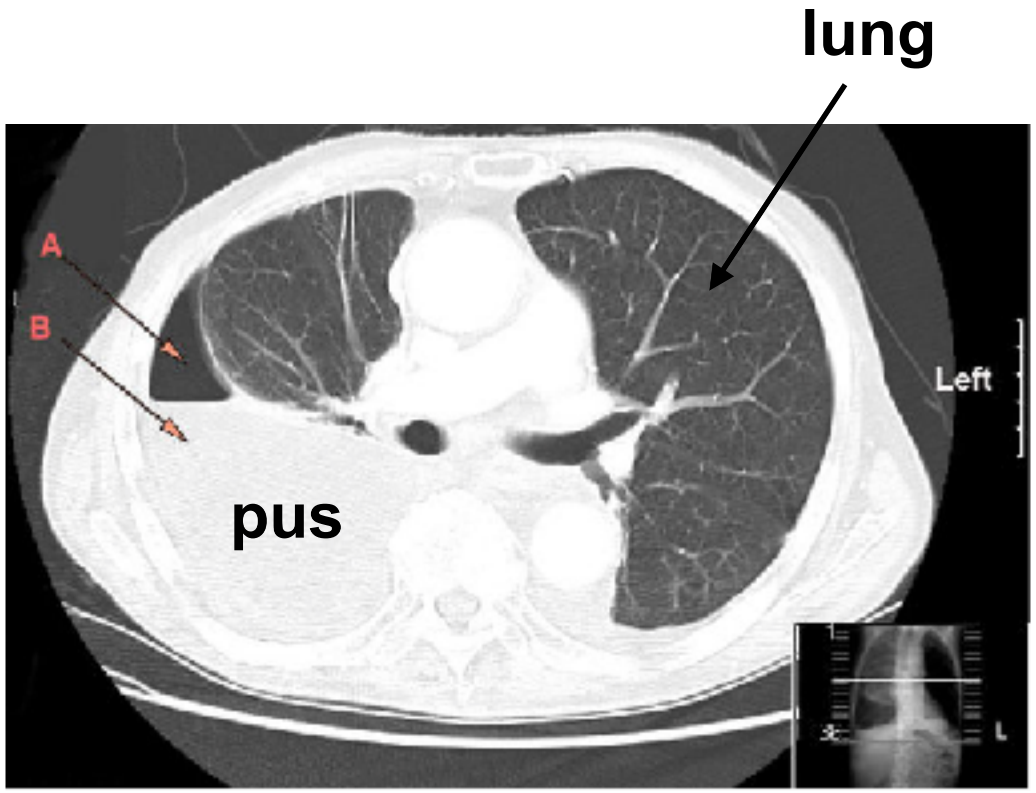
of bacterial infectious disease

Pyothorax

Chest infection
Accumulation of pus in chest cavity
Secondary to pneumonia
>65 years old
Immediate antibiotics recommended
Bacterial culture positivity: ~50%
Culture takes days

Known causative pathogens are:

S. pneumoniae,
S aureus,
Anaerobic bacteria,
(Fusobacterium, Prevotella,
Peptostreptococcus, Bacteroides)
etc

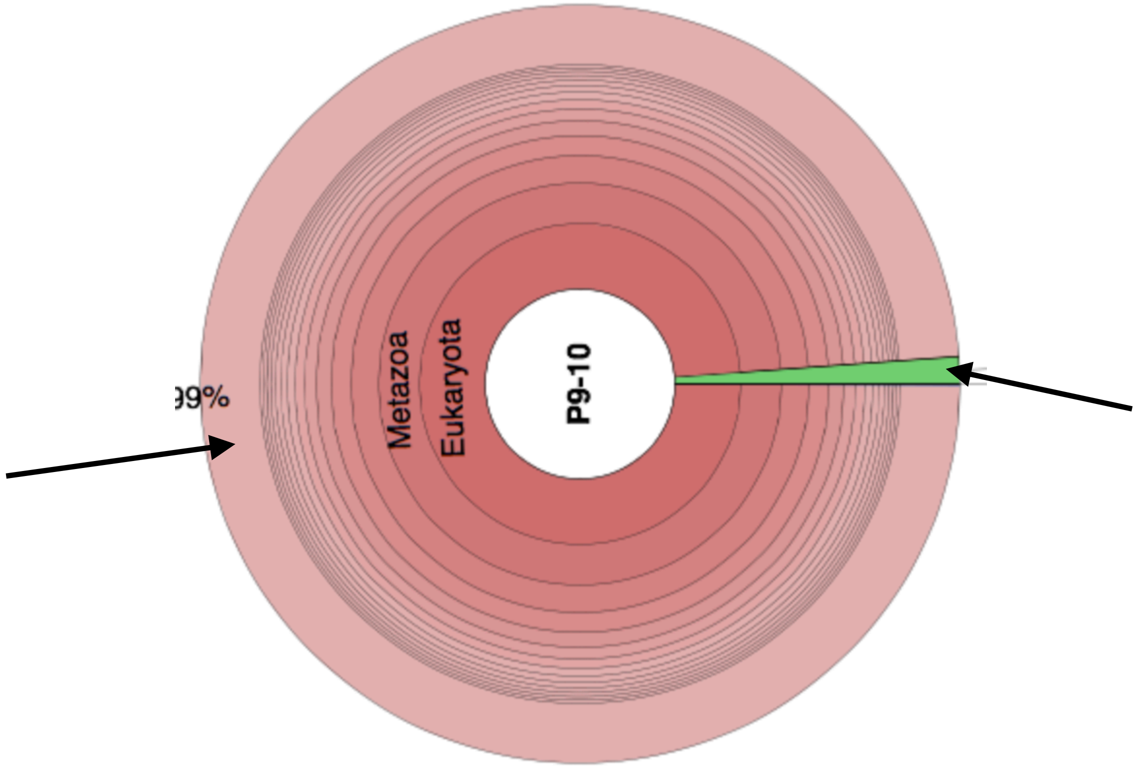


Wikipedia

host DNA is abundunt in clinical samples

Pleural effusion from pyothorax patient

Bacterial DNA / host DNA amount : 0.03~0.3%



Human

30M reads

Bacteria

9000 reads

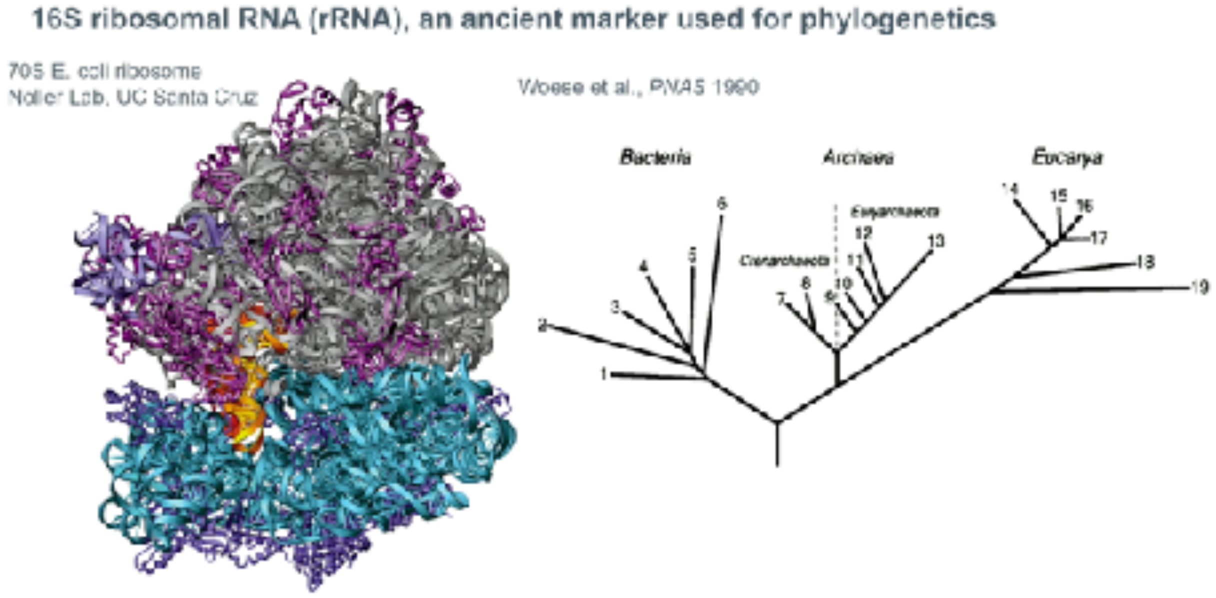
One MiSeq run ~ €2000 / sample

16S ribosomal DNA sequencing

(Nakagawa et al. 2010. PNAS)

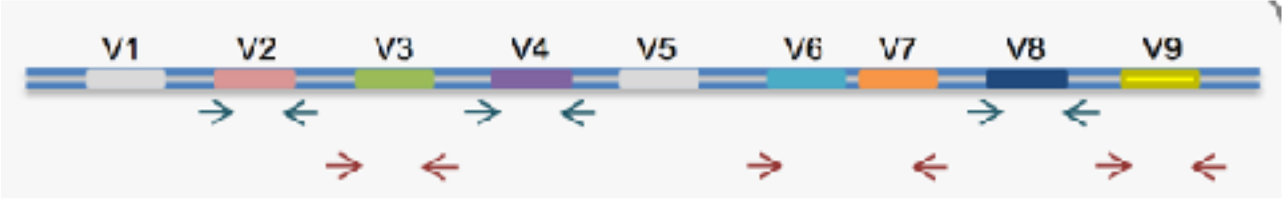
Universal primers

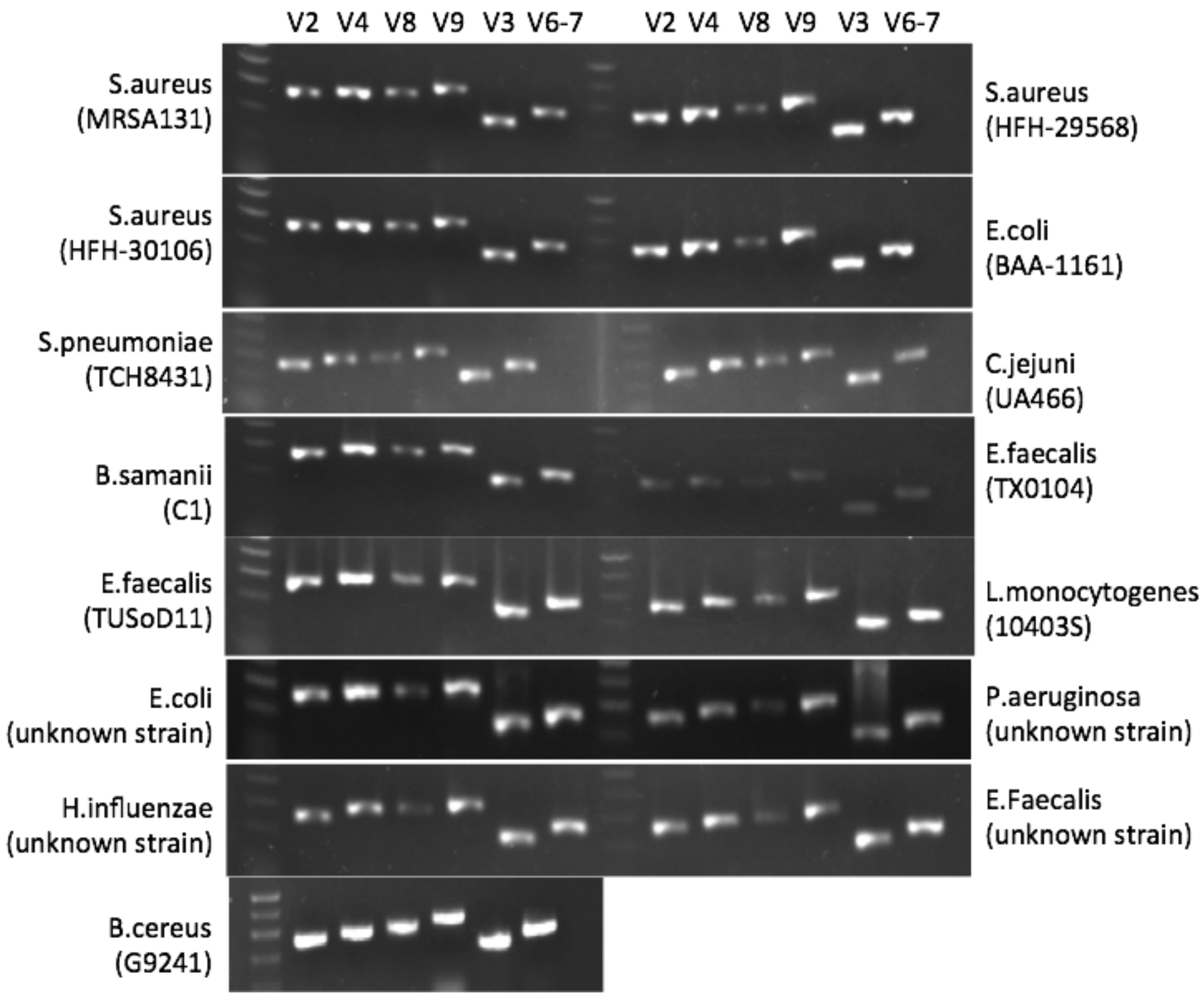
Next gen sequencing



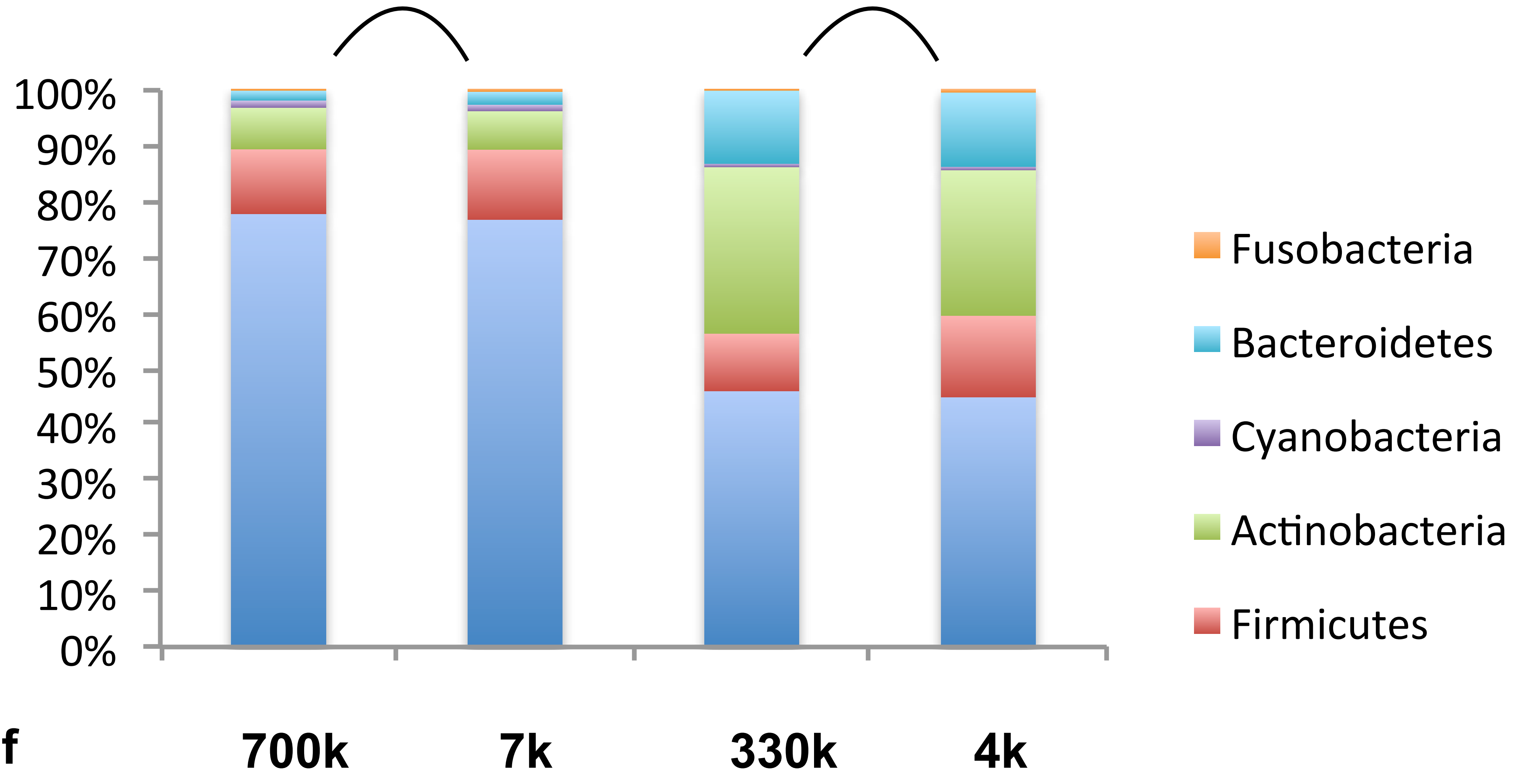
Ion PGM sequencer for 16S sequencing

16S metagenomics kit





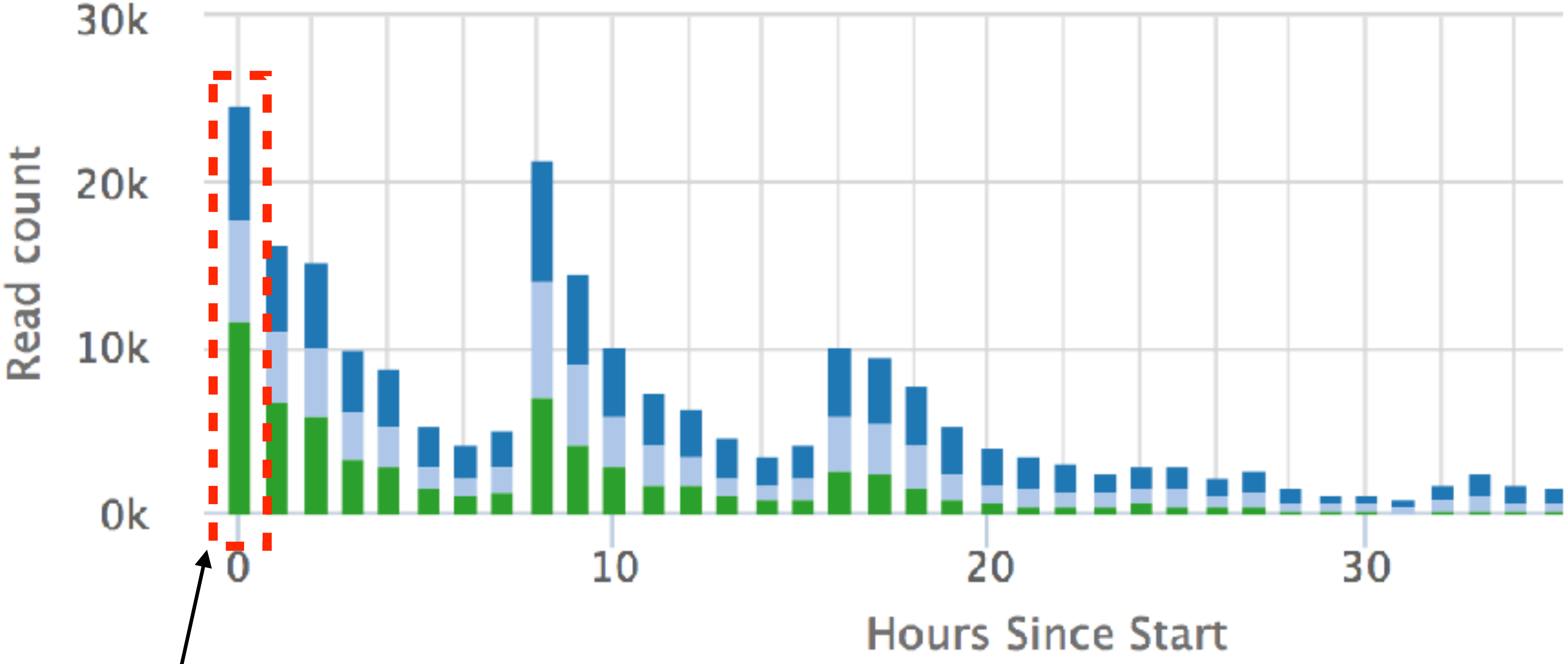
16S seq results with different read counts in IonPGM



number of sequence reads

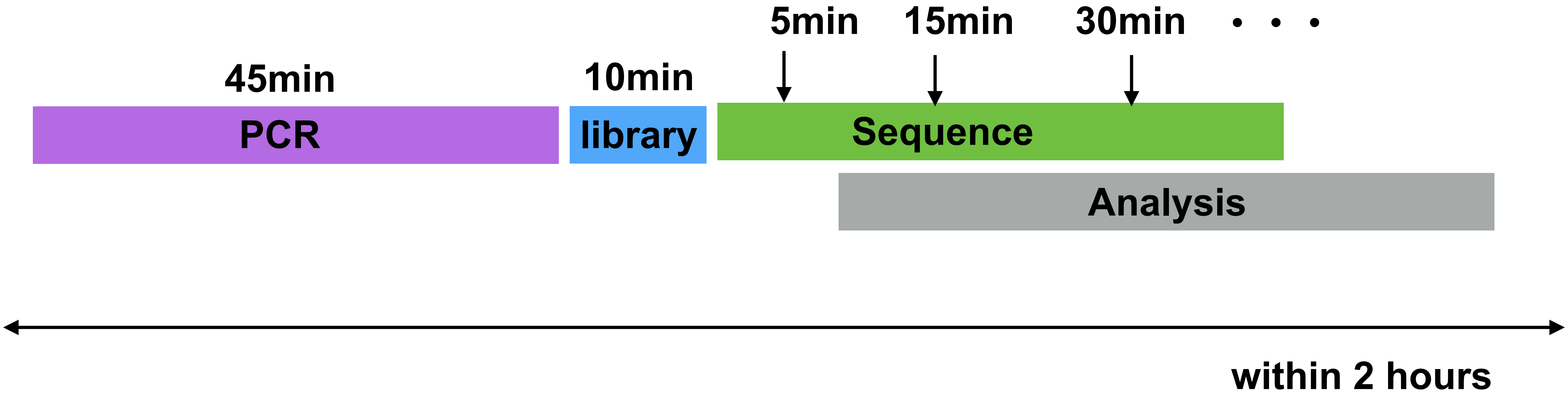
How many reads are necessary before diagnosis?

Time course for amplicon sequence



First 1 hour = 24k reads

Maybe first few minutes sequencing is enough



5min 15min 30min • • •

45min 10min

PCR

library

Sequence

Analysis

within 2 hours

Rapid sequencing kit (SQK-RAD002)

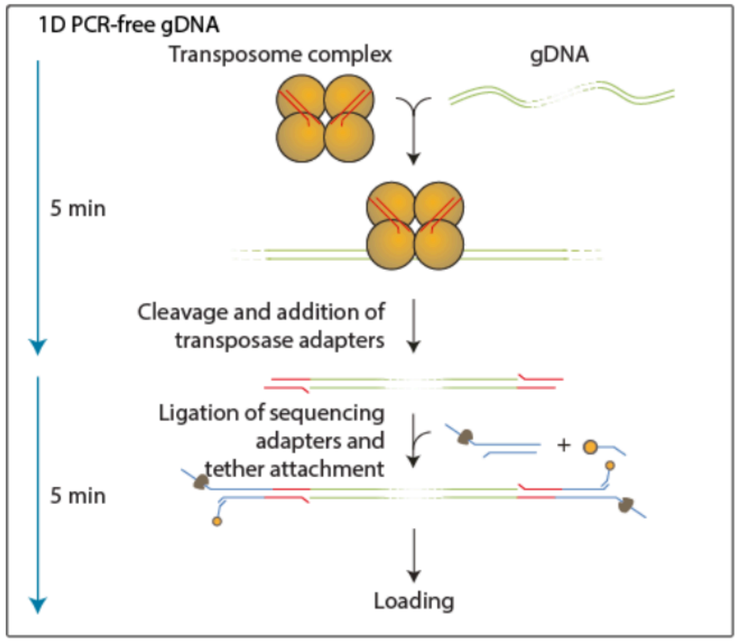


Figure RSWF: Workflow of the Rapid Sequencing library preparation protocol.

Control bacterial mock community with equimolar 16S

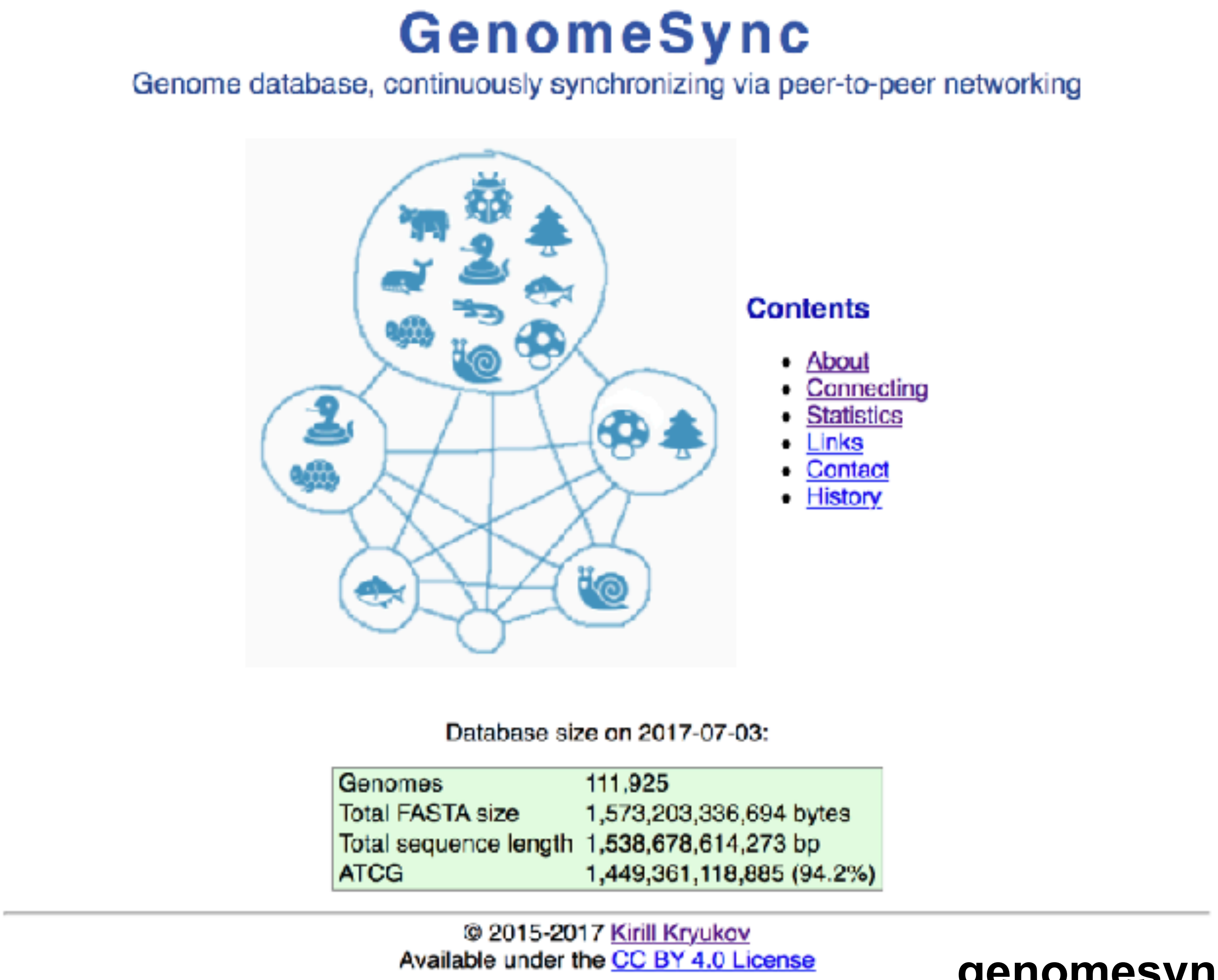
Short read sequence
(IonPGM)

Long read sequence
(MinION)

	Species	strain	Genome size	16S copy number
1	*Acinetobacter baumannii*	strain 5377	3,976,747	6
2	*Actinomyces odontolyticus*	strain 1A.21	2,391,230	3
3	*Bacillus cereus*	strain NRS 248	5,224,283	13
4	*Bacteroides vulgatus*	strain ATCC® 8482ä	5,163,189	7
5	*Clostridium beijerinckii*	strain NCIMB 8052	6,000,632	14
6	*Deinococcus radiodurans*	strain R1 (smooth)	2,648,638	2
7	*Enterococcus faecalis*	strain OG1RF	2,739,625	4
8	*Escherichia coli*	strain MG1655	4,641,652	7
9	*Helicobacter pylori*	strain 26695	1,667,867	2
10	*Lactobacillus gasseri*	strain 63 AM	1,894,360	6
11	*Listeria monocytogenes*	strain EGDe	2,944,528	6
12	*Neisseria meningitidis*	strain MC58	2,272,360	4
13	*Propionibacterium acnes*	strain KPA171202	2,560,265	3
14	*Pseudomonas aeruginosa*	strain PAO1-LAC	6,264,404	4
15	*Rhodobacter sphaeroides*	strain ATH 2.4.1	4,131,542	4
16	*Staphylococcus aureus*	strain TCH959	2,830,017	5
17	*Staphylococcus epidermidis*	strain PCI 1200	2,499,279	6
18	*Streptococcus agalactiae*	strain 2603 V/R	2,160,267	7
19	*Streptococcus mutans*	strain UA159	2,032,925	5
20	*Streptococcus pneumoniae*	strain TIGR4	2,160,842	4

BEI resource
(HM-782D)

Analysis 1: Blast search using GenomeSync database



GenomeSync

Genome database, continuously synchronizing via peer-to-peer networking

Contents

- About
- Connecting
- Statistics
- Links
- Contact
- History

Database size on 2017-07-03:

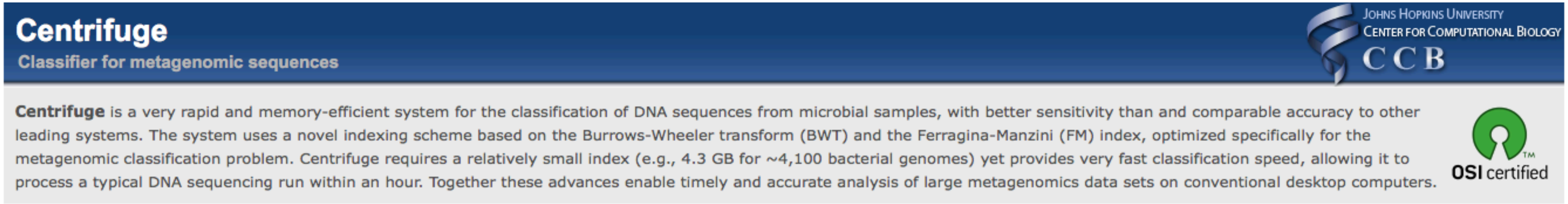
Genomes	111,925
Total FASTA size	1,573,203,336,694 bytes
Total sequence length	1,538,678,614,273 bp
ATCG	1,449,361,118,885 (94.2%)

Kirill Kryukov

genomesync.org

> Genome Search Took Kit (GSTK)

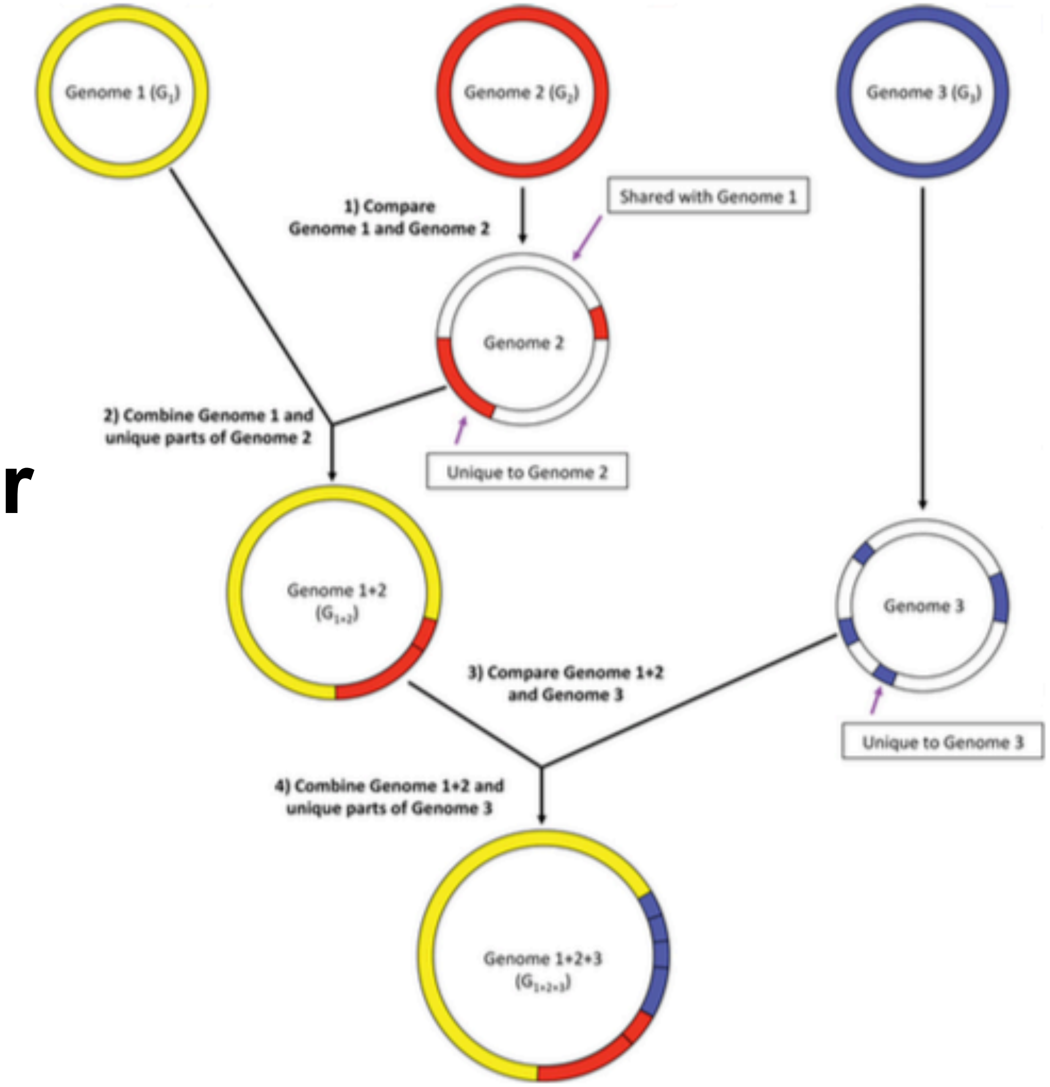
Analysis 2: Centrifuge



Centrifuge

Classifier for metagenomic sequences

Centrifuge is a very rapid and memory-efficient system for the classification of DNA sequences from microbial samples, with better sensitivity than and comparable accuracy to other leading systems. The system uses a novel indexing scheme based on the Burrows-Wheeler transform (BWT) and the Ferragina-Manzini (FM) index, optimized specifically for the metagenomic classification problem. Centrifuge requires a relatively small index (e.g., 4.3 GB for ~4,100 bacterial genomes) yet provides very fast classification speed, allowing it to process a typical DNA sequencing run within an hour. Together these advances enable timely and accurate analysis of large metagenomics data sets on conventional desktop computers.



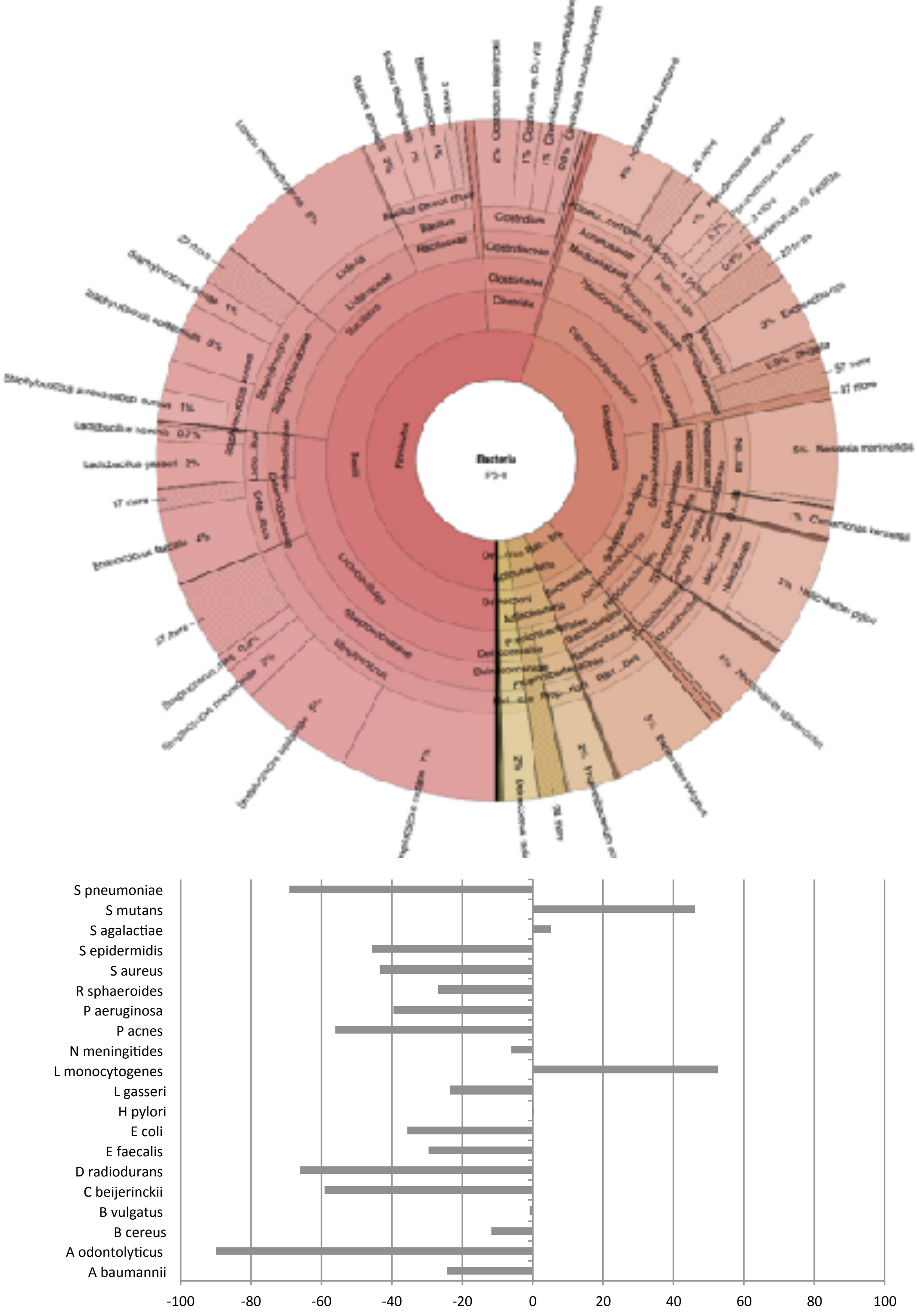
Rapid metagenomic sequence classifier

~4100 bacterial genomes

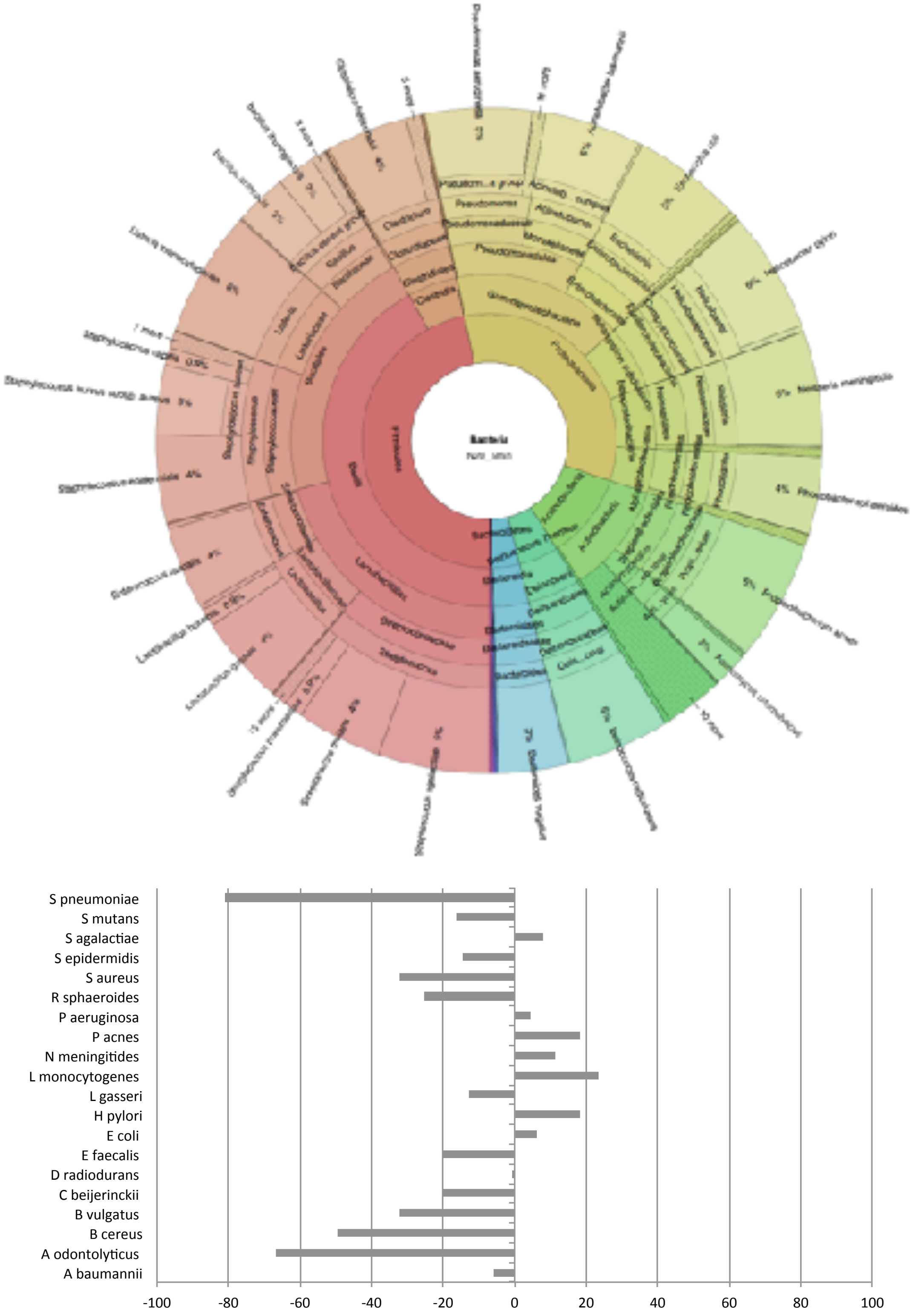
(Kim *et al.* 2016 Genome Research)

Spesies level classification

Short read (IonPGM)

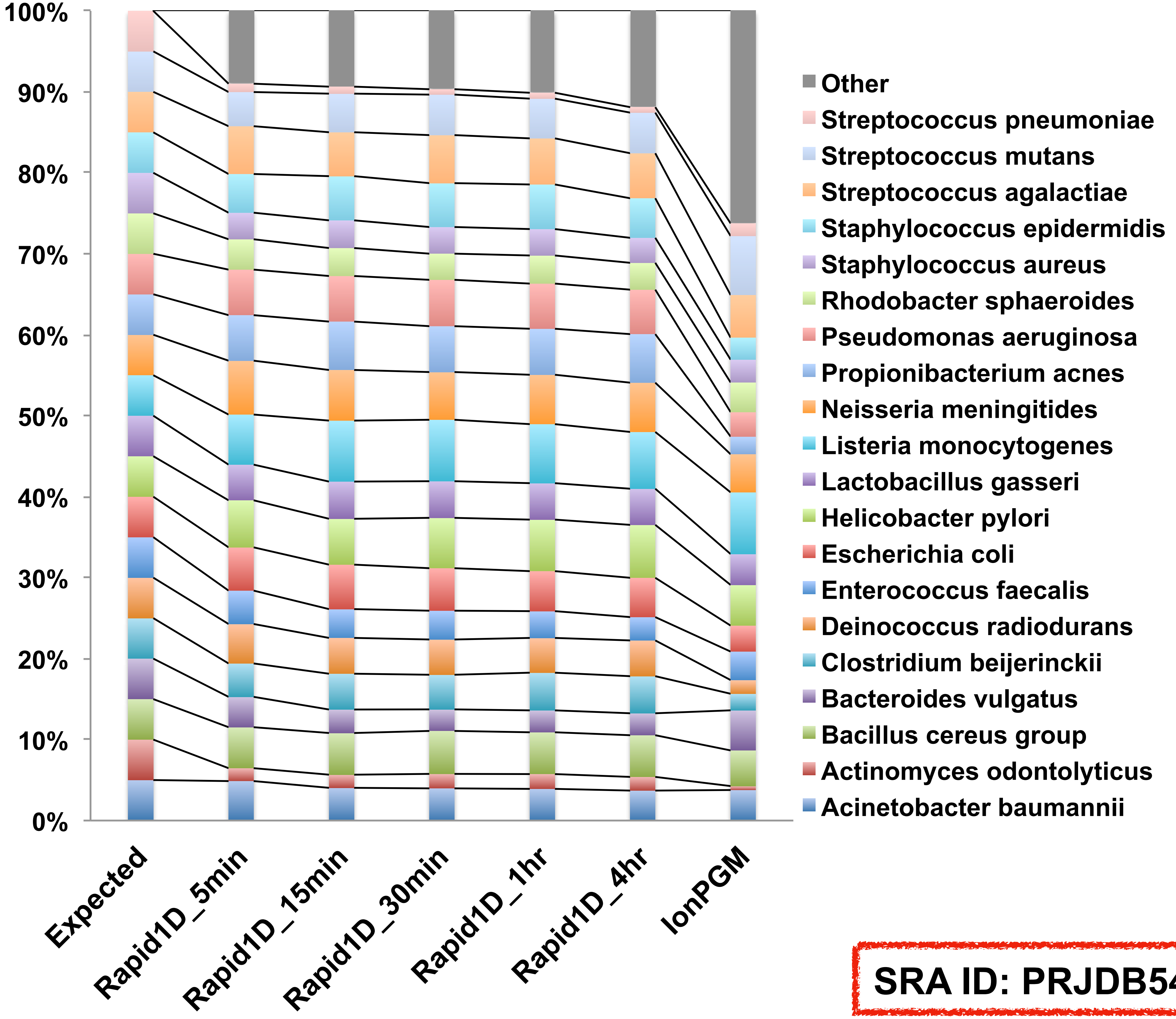


Long read (MinION)



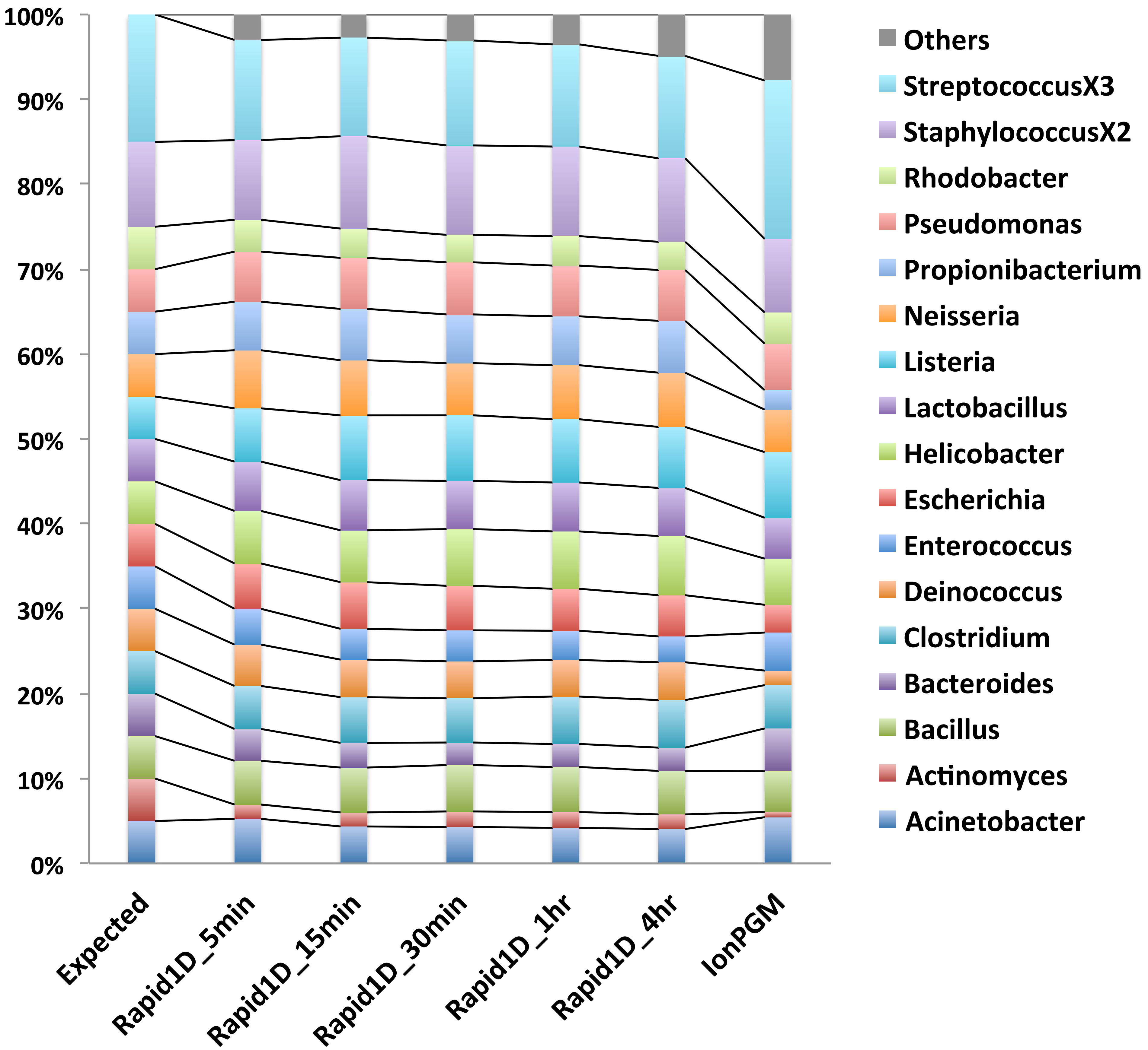
(Mitsuhashi, et al. 2017 Sci Rep)

Result of analysis 1: BLAST/GenomeSync (species level)

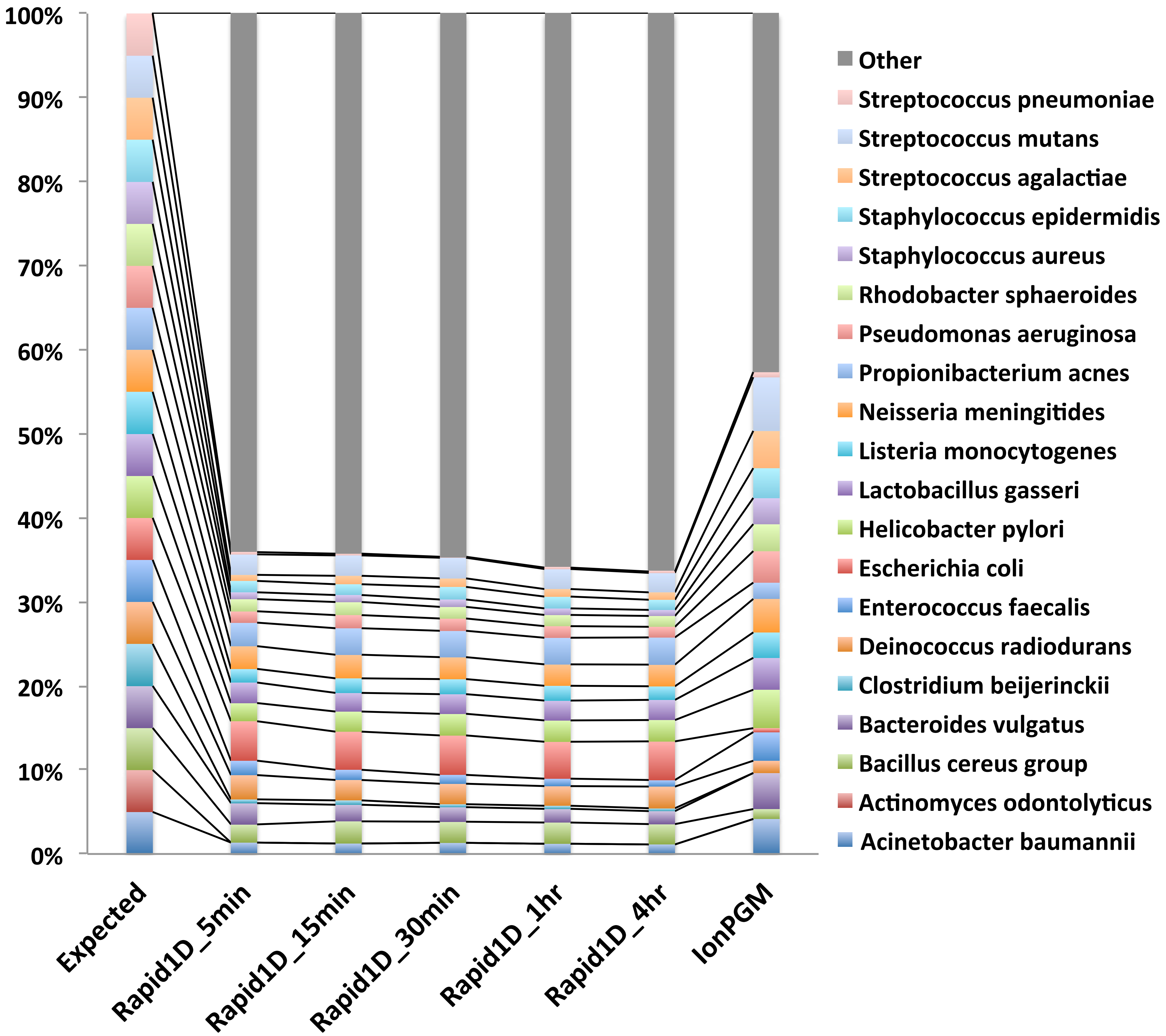


SRA ID: PRJDB5408

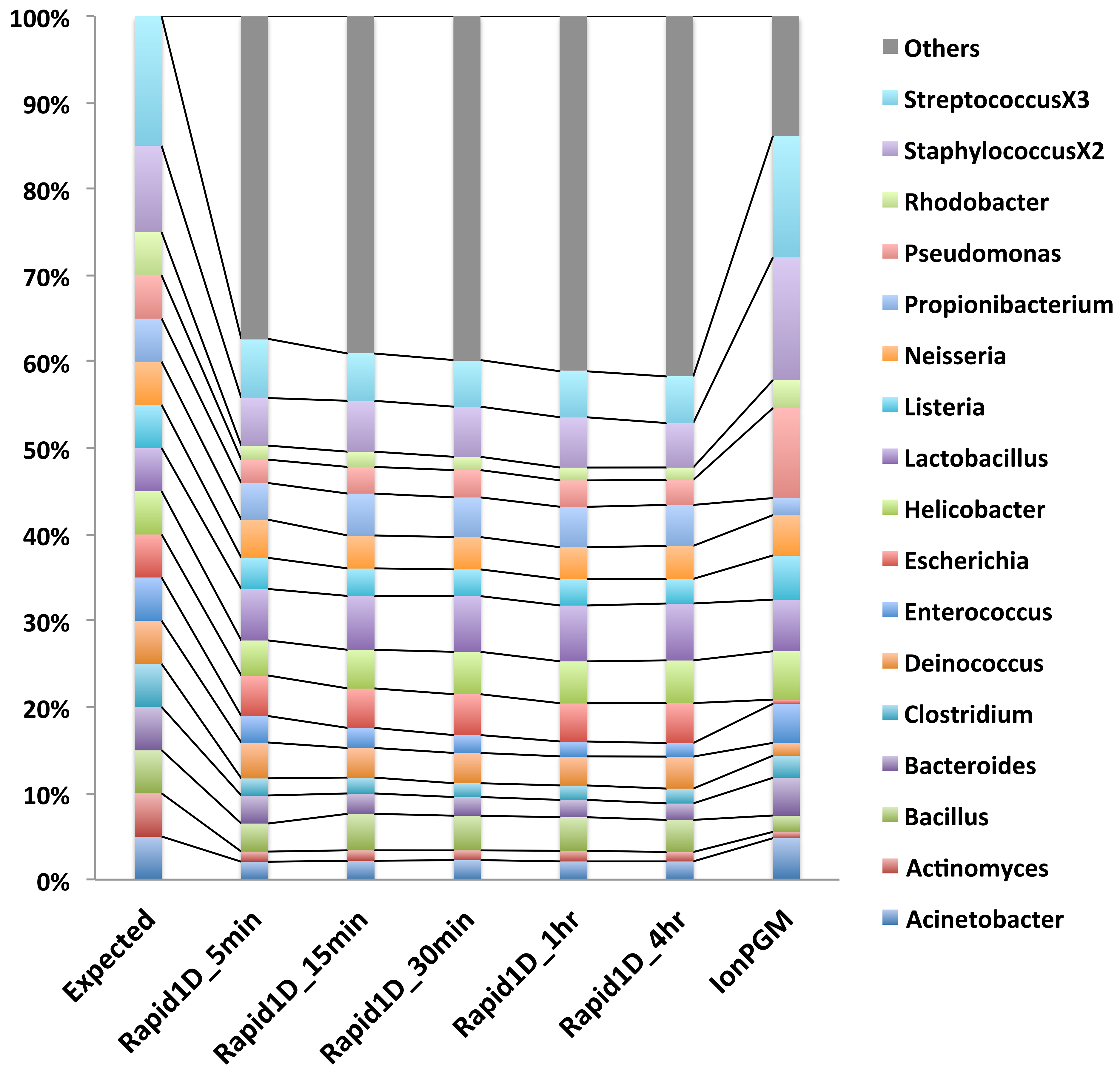
Result of analysis 1: BLAST/GenomeSync (genus level)



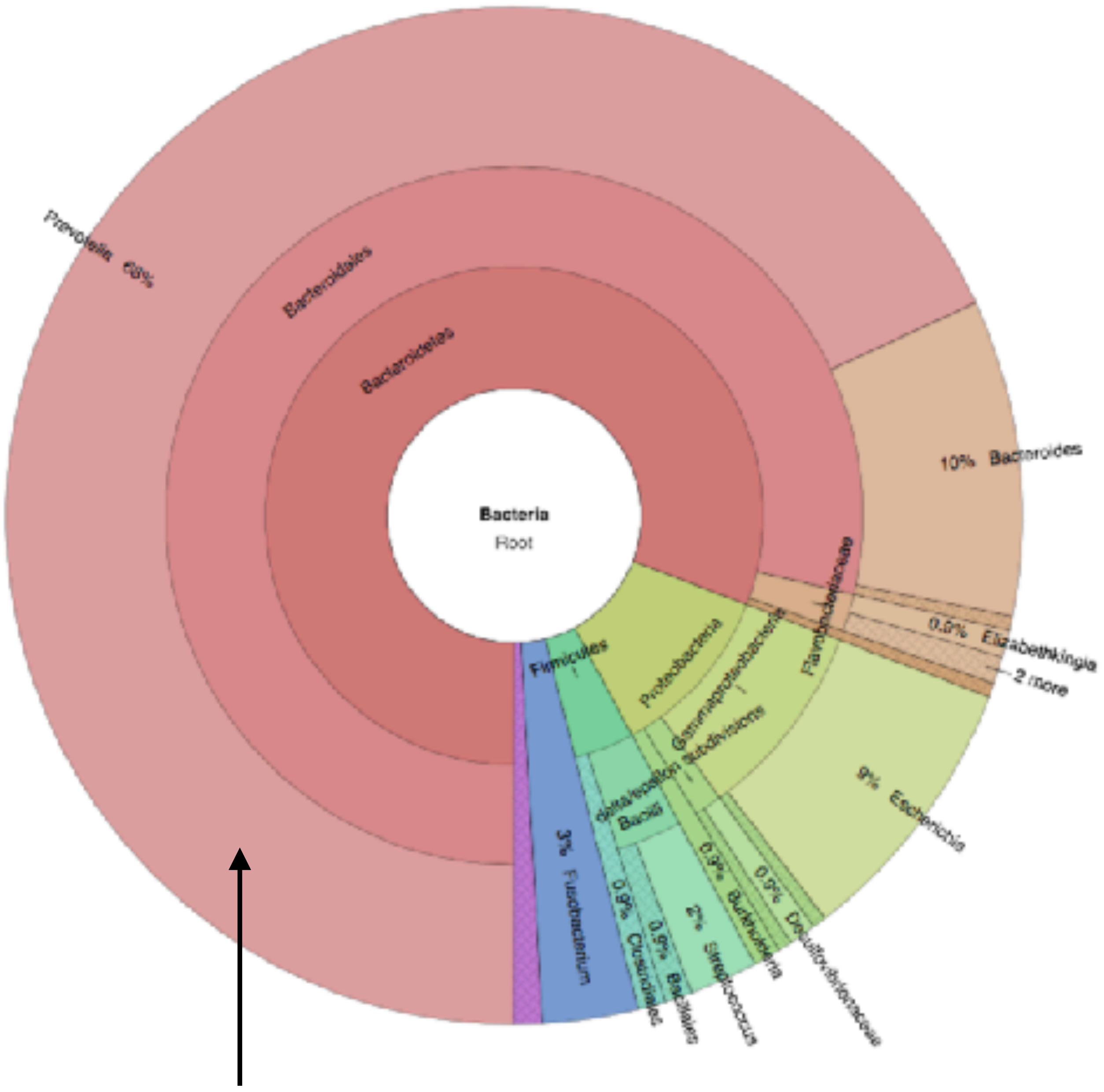
Result of analysis 2: Centrifuge (species level)



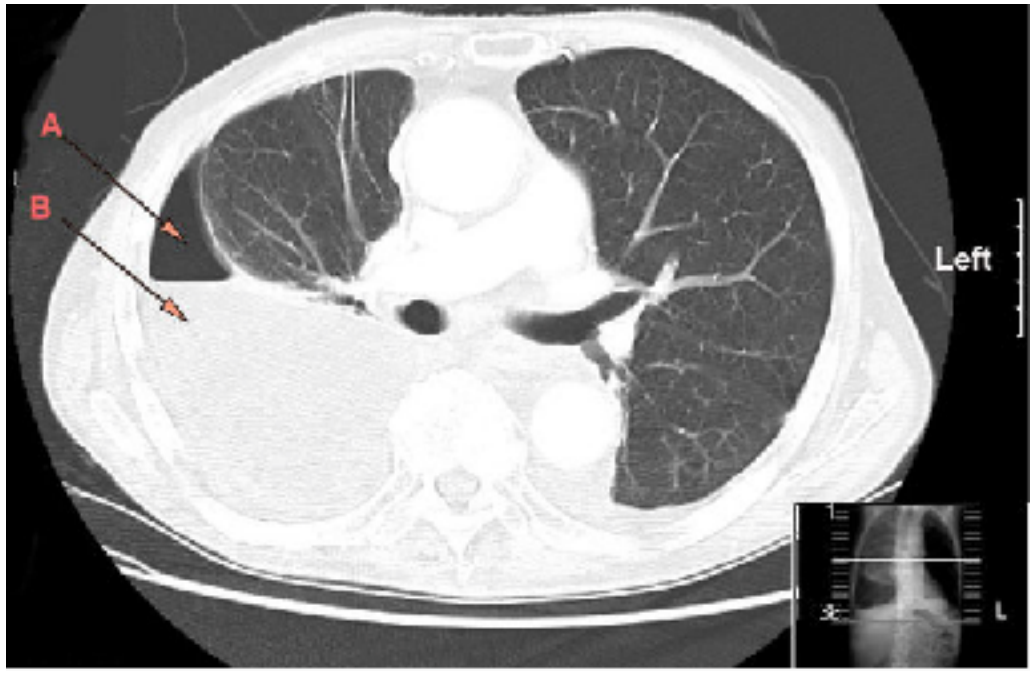
Result of analysis 2: Centrifuge (genus level)



5-min sequence of pleural effusion from pyothorax patient



Prevotella oris was detected



The biggest barrier for clinical use

Metagenomic (MiSeq)	~ €2000 / sample
16S sequencing (IonPGM)	~ €200 / sample
MinION 16S sequencing?	~ €700 / flow cell

¥¥

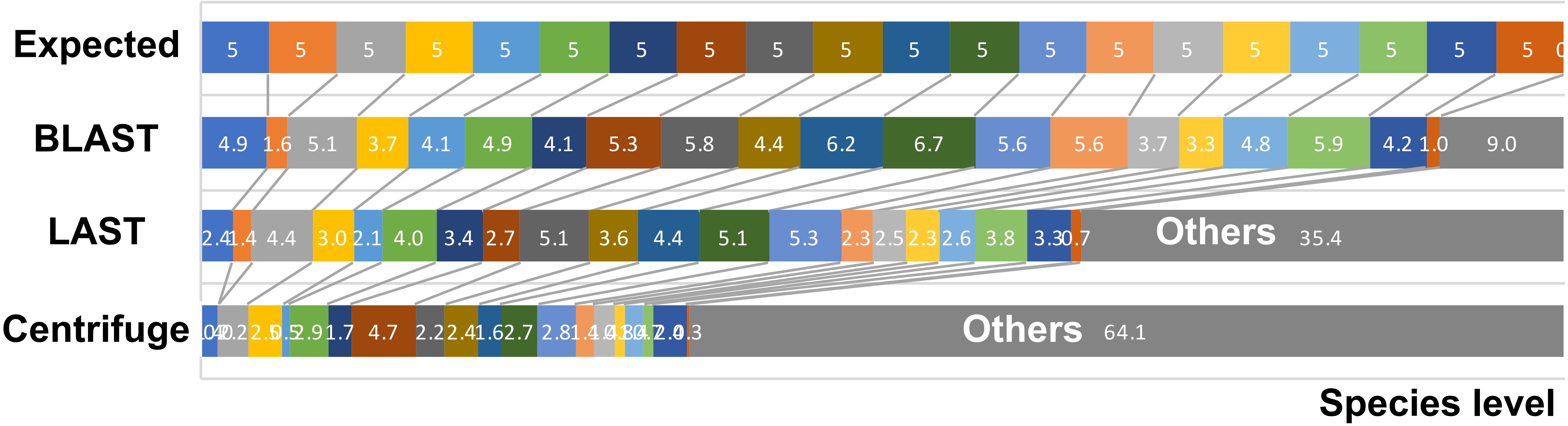
Bacterial culture	~ €15 / sample

price will decrease?

FLONGLE
(“flow cell dongle”) → **“MinION Dx”**

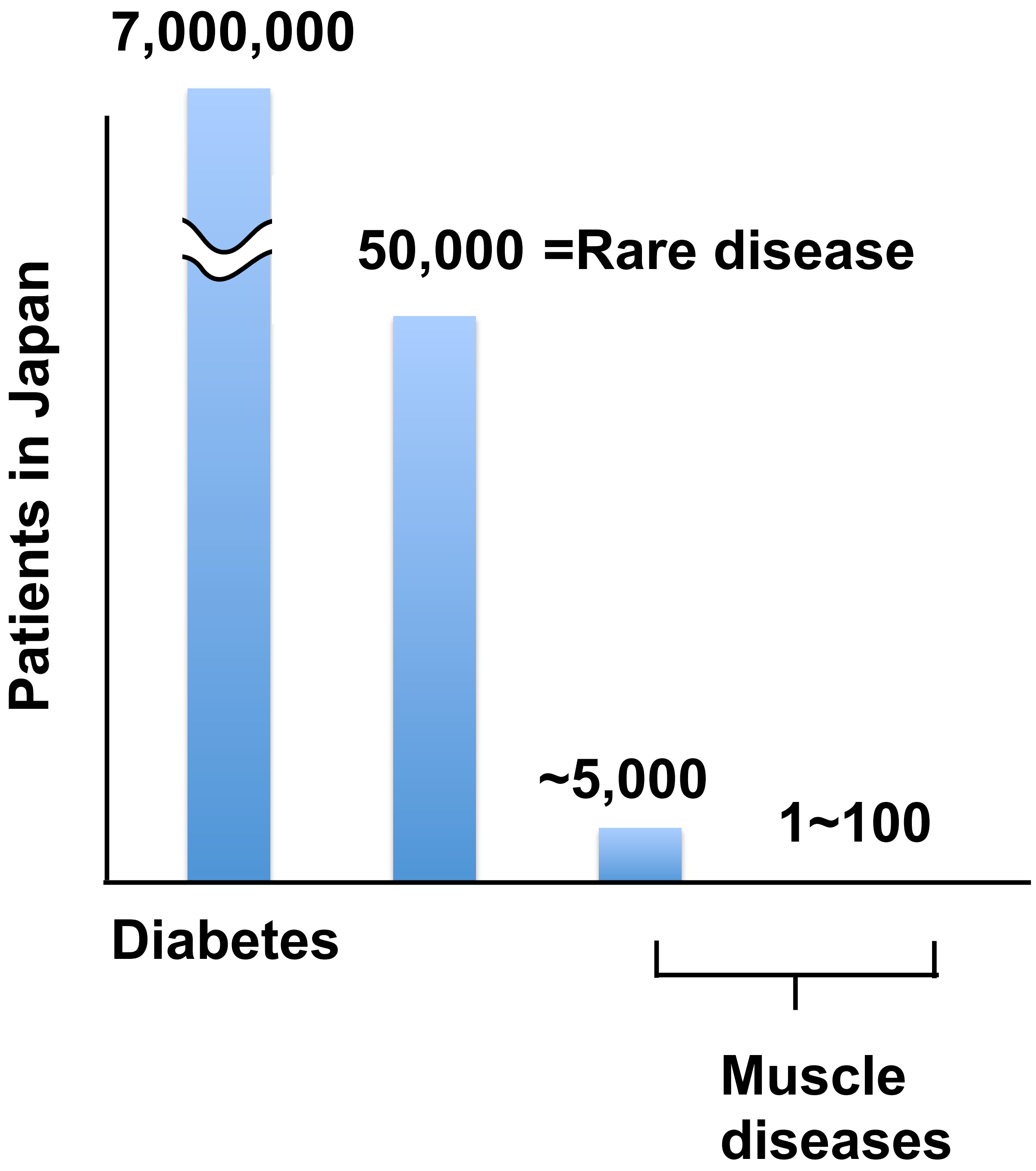
Development of a new pipeline using LAST

- **LAST(Kiełbasa et al. *Genome Res*. 2011)**
- **Speed: Centrifuge > LAST >>> BLAST**
- **Accuracy：Centrifuge < LAST << BLAST**



→ D4Z4 Repeat sequencing for the diagnosis of a muscular dystrophy

Muscular dystrophies



from the book of Dr. Ikuya Nonaka “Rinsho no tameno kin-byori”

Facioscapulohumeral muscular dystrophy (FSHD)

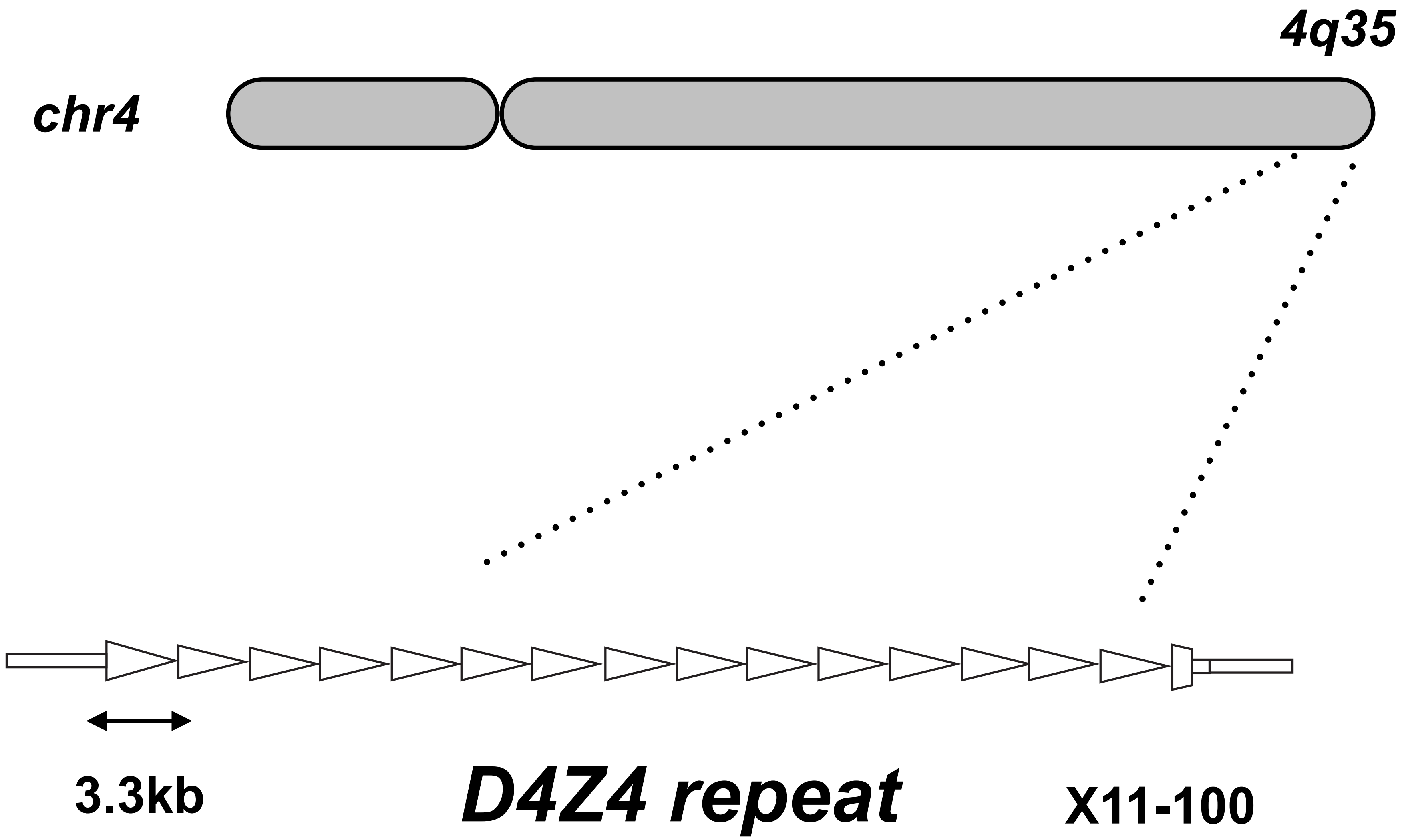
FSH Society
https://www.fshsociety.org/

Autosomal dominant muscular dystrophy

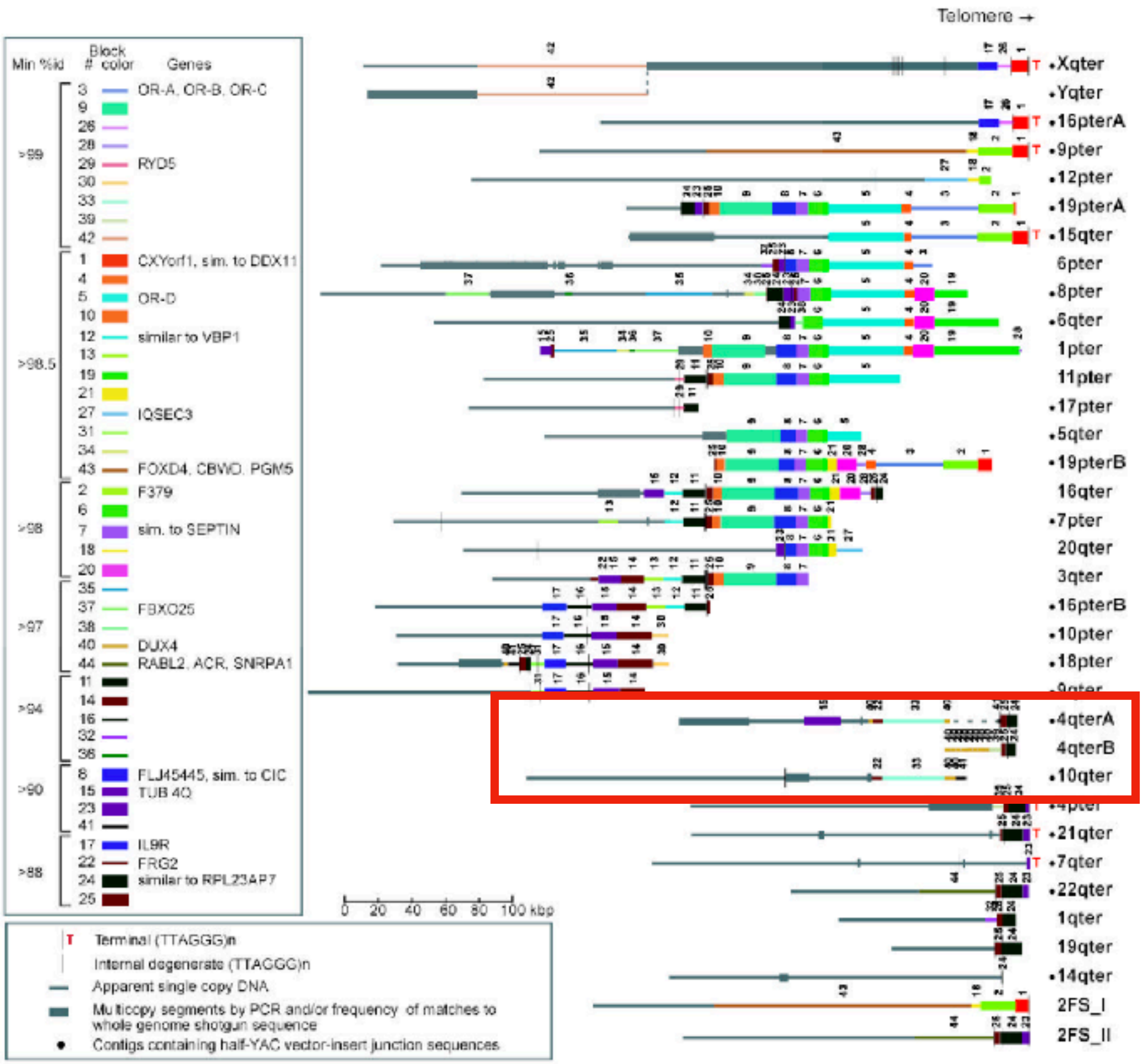
Facial, scapular and humeral muscle weakness

Prevalence: estimated ~1/10,000

D4Z4 subtelomeric macrosatellite repeat is linked to FSHD



Subtelomeric repeat sequences

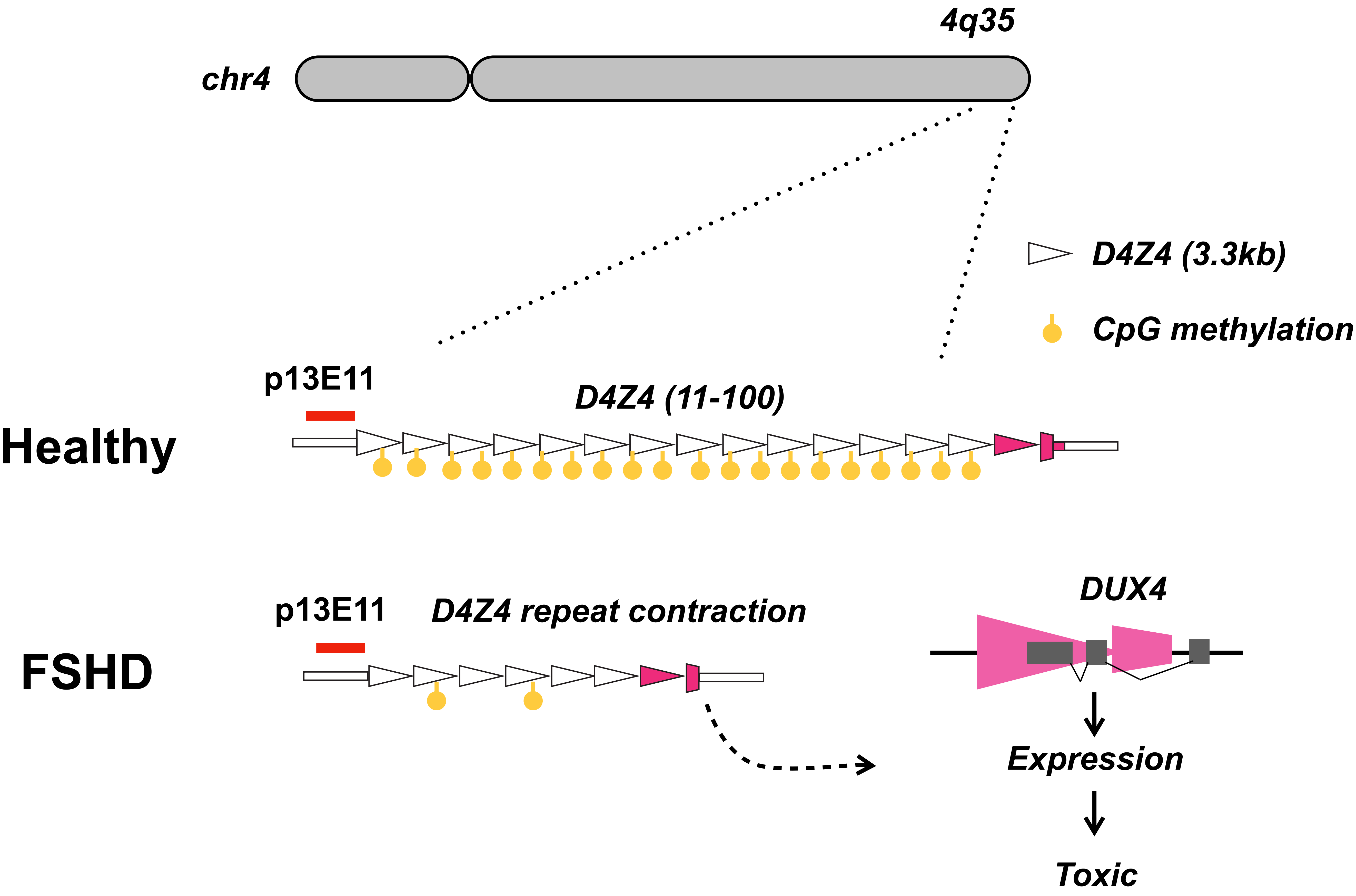


Variations!

D4Z4

(Linardopoulou et al. Nature 2005)

Contraction of D4Z4 array causes FSHD



4q35

chr4

▷ *D4Z4 (3.3kb)*

● *CpG methylation*

Healthy

p13E11

D4Z4 (11-100)

FSHD

p13E11

D4Z4 repeat contraction

DUX4

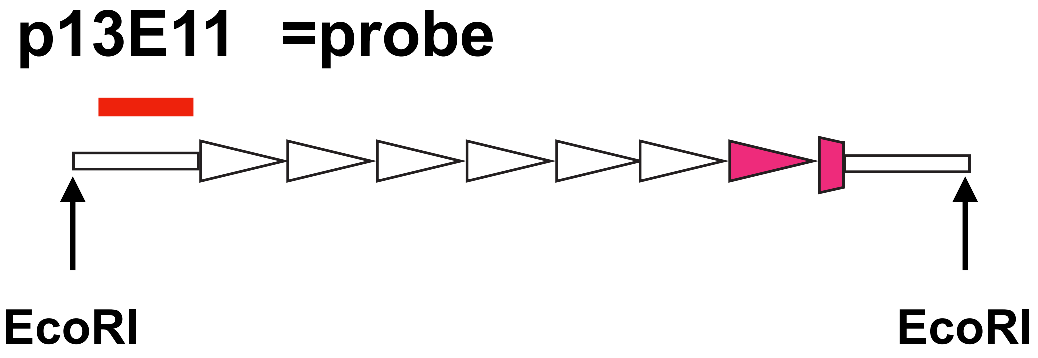
↓

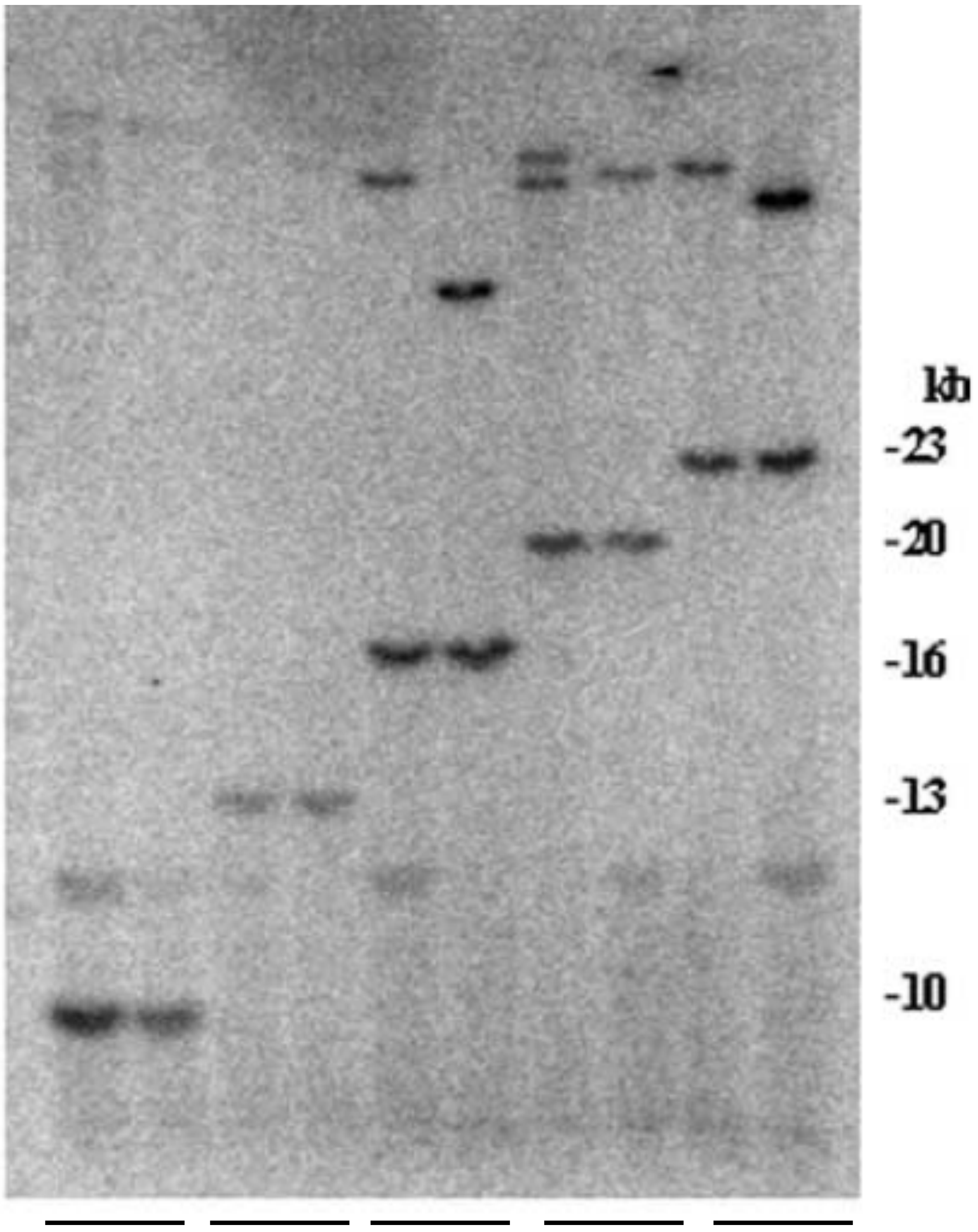
Expression

↓

Toxic

Southern blot analysis for D4Z4 repeat size



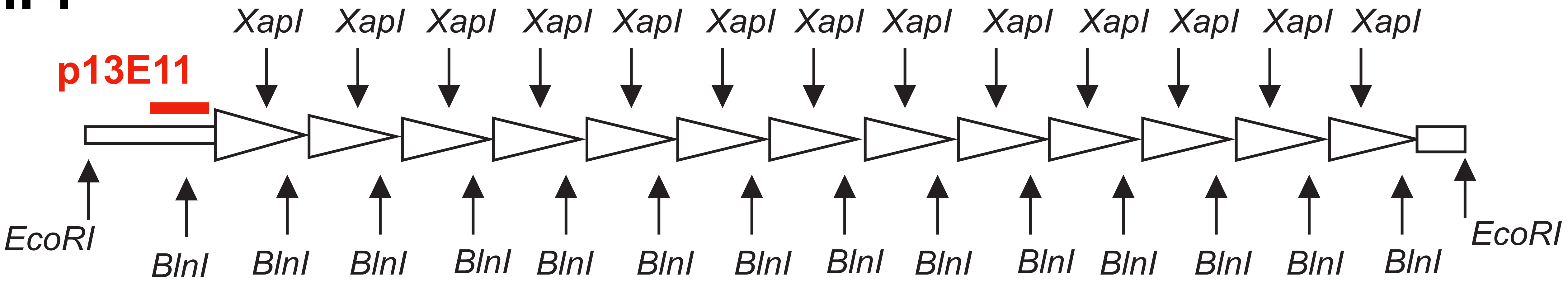


# of D4Z4 repeat: 1 2 3 4 5

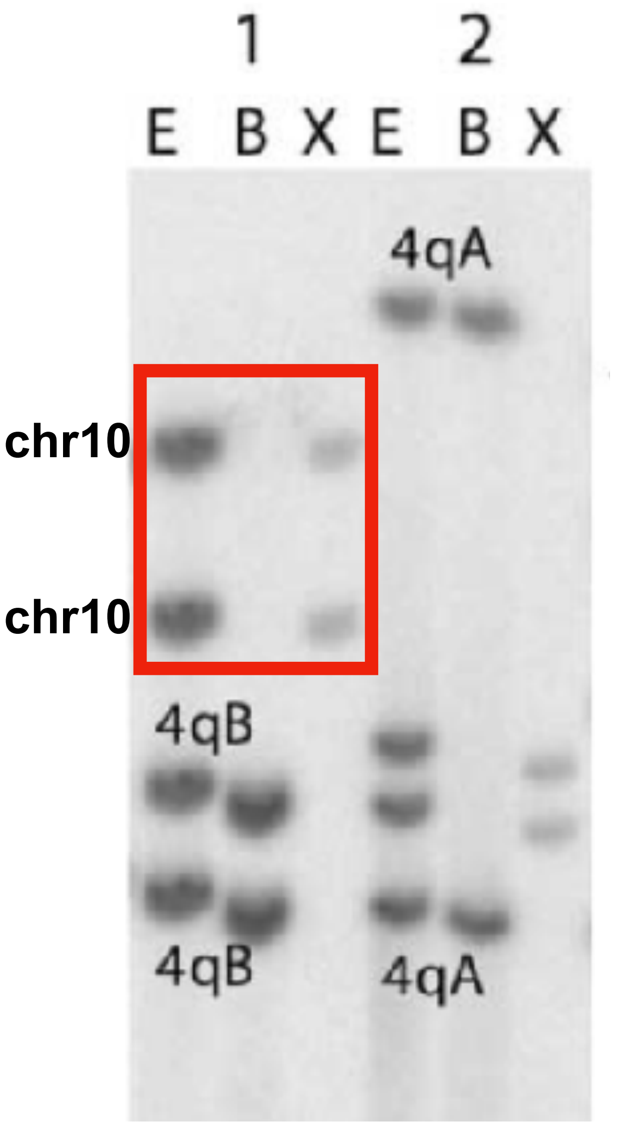
(Goto *et al.* 2006 Neuromuscular Disorders)

D4Z4 homologous repeats on chr4 and chr10

chr4



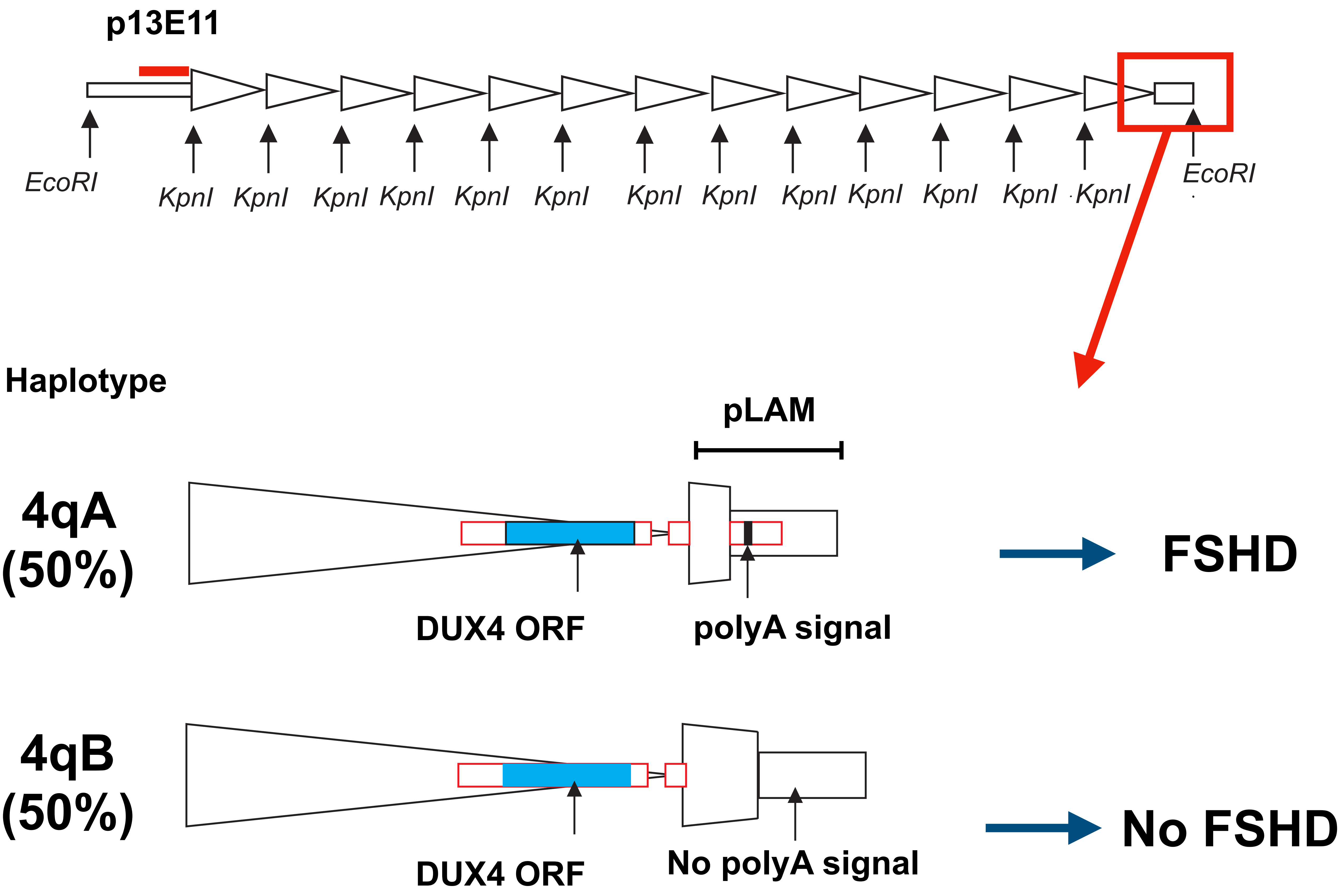
chr10



←**Restriction enzyme digestion**

(Lemmers *et al.* 2007)

There are two haplotypes in D4Z4



How to make a genetic diagnose for FSHD

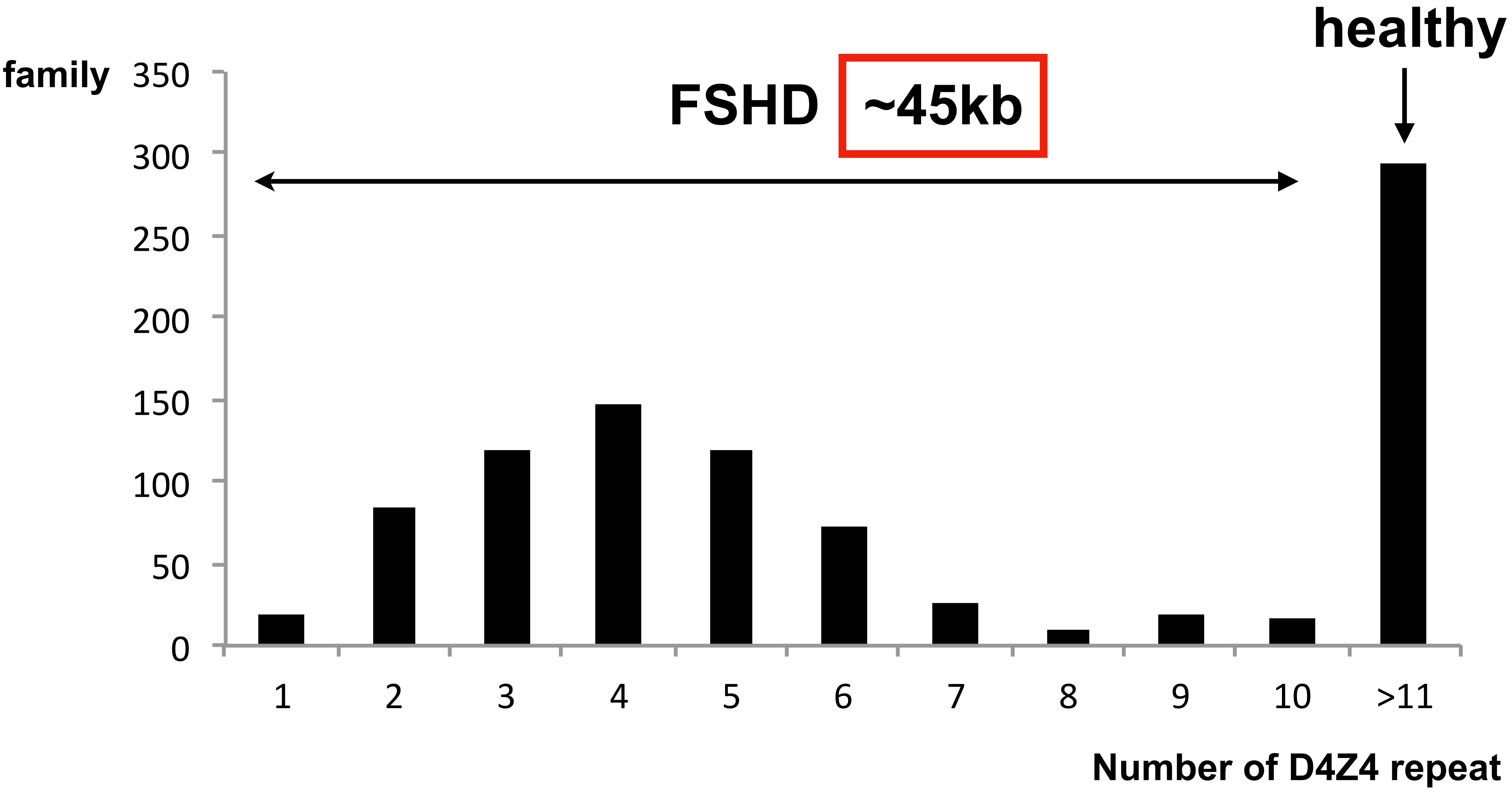
1. number of repeat	Southernblotting
2. distinguish homologous repeat	Southernblotting
3. haplotype (4qA or 4qB)	Southernblotting

😩

↓

Sequencing D4Z4 repeat using MinION

Distribution of repeat number in FSHD patients



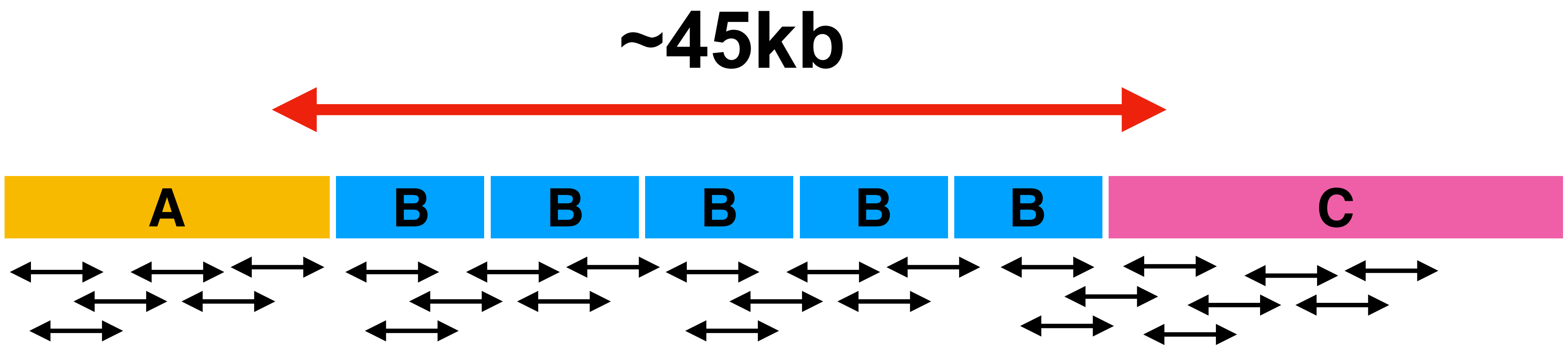
(unpublished data)

Golden Rule

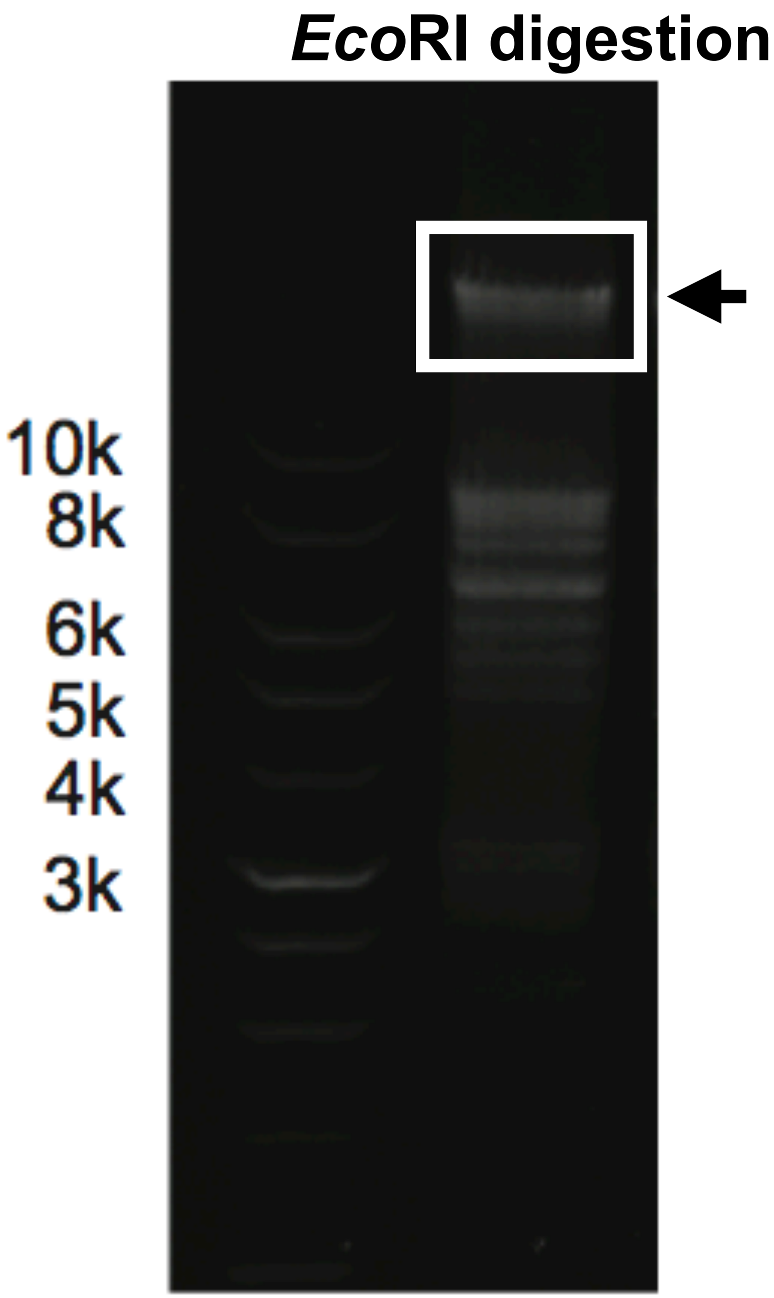
In order to read repeat sequence

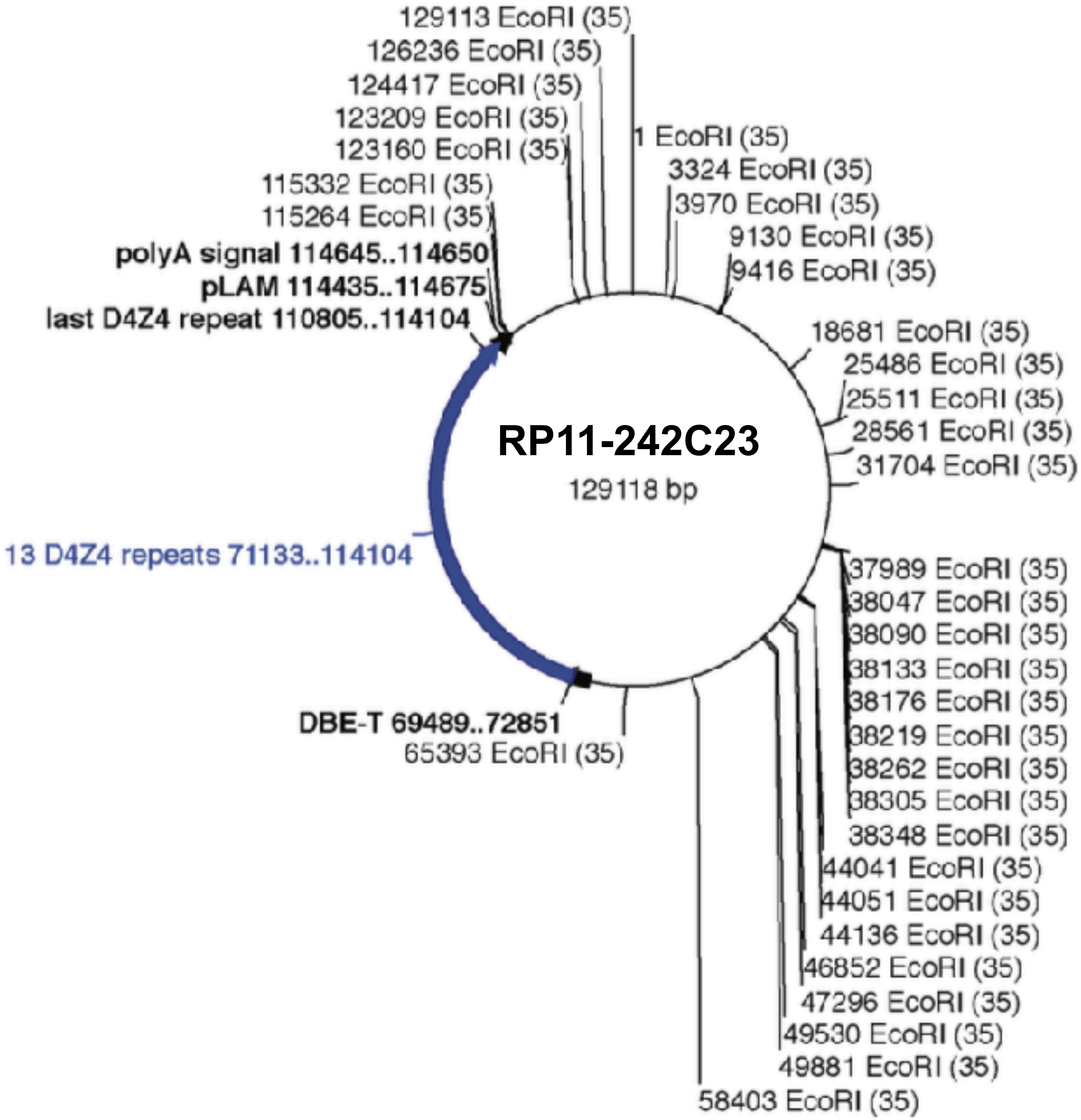
you should read

LONGER THAN THE REPEAT



BAC clone containing 13 D4Z4 repeats





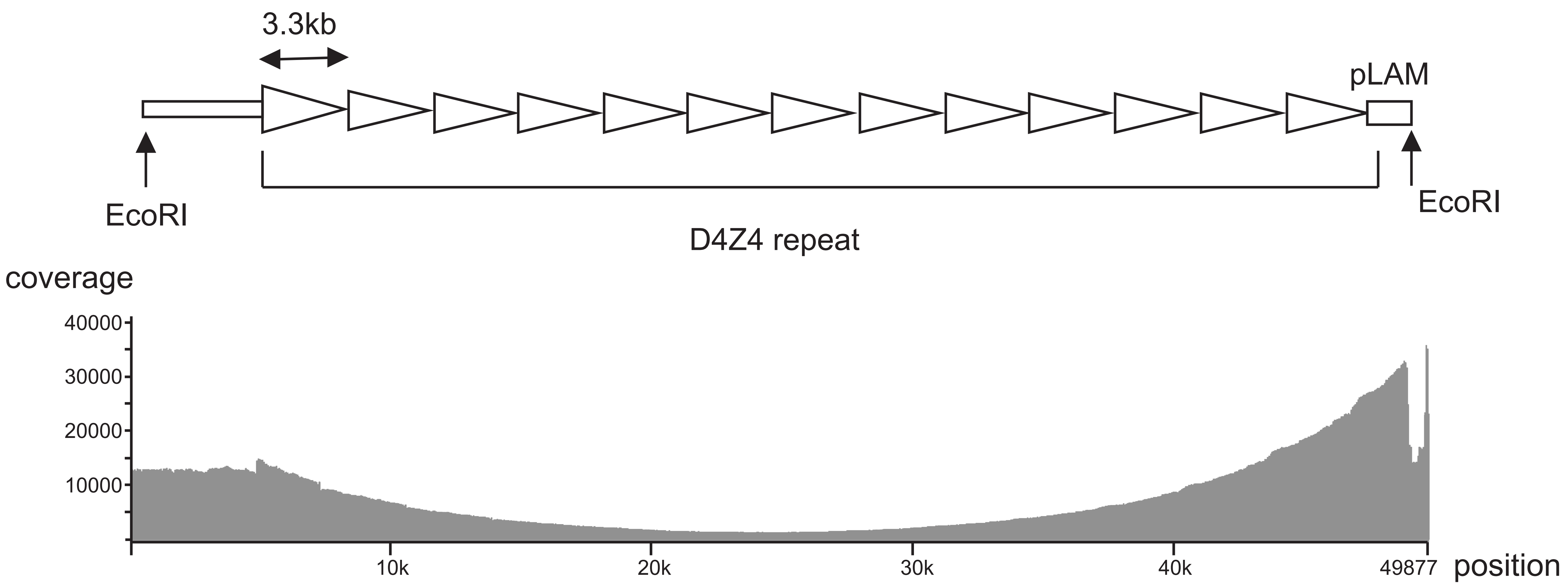
method

BAC clone: RP11-242C23

enzyme digestion	***Eco*RI digestion + Klenow** **gel extraction**
	↓
library preparation	**SQK-LSK108 (1D sequencing kit)**
	↓
sequencing	**MinION Mk1B** **R9.4 flowcell**
	↓
base-calling	**Albacore v.1.1.0**
	↓
alignment	***Eco*RI fragment (Genbank: CT476828.7)** **lastal | last-split**
	↓
genotyping	**last-genotype** **/most common nucleotide**

Aligning reads to D4Z4 reference using LAST

http://last.cbrc.jp

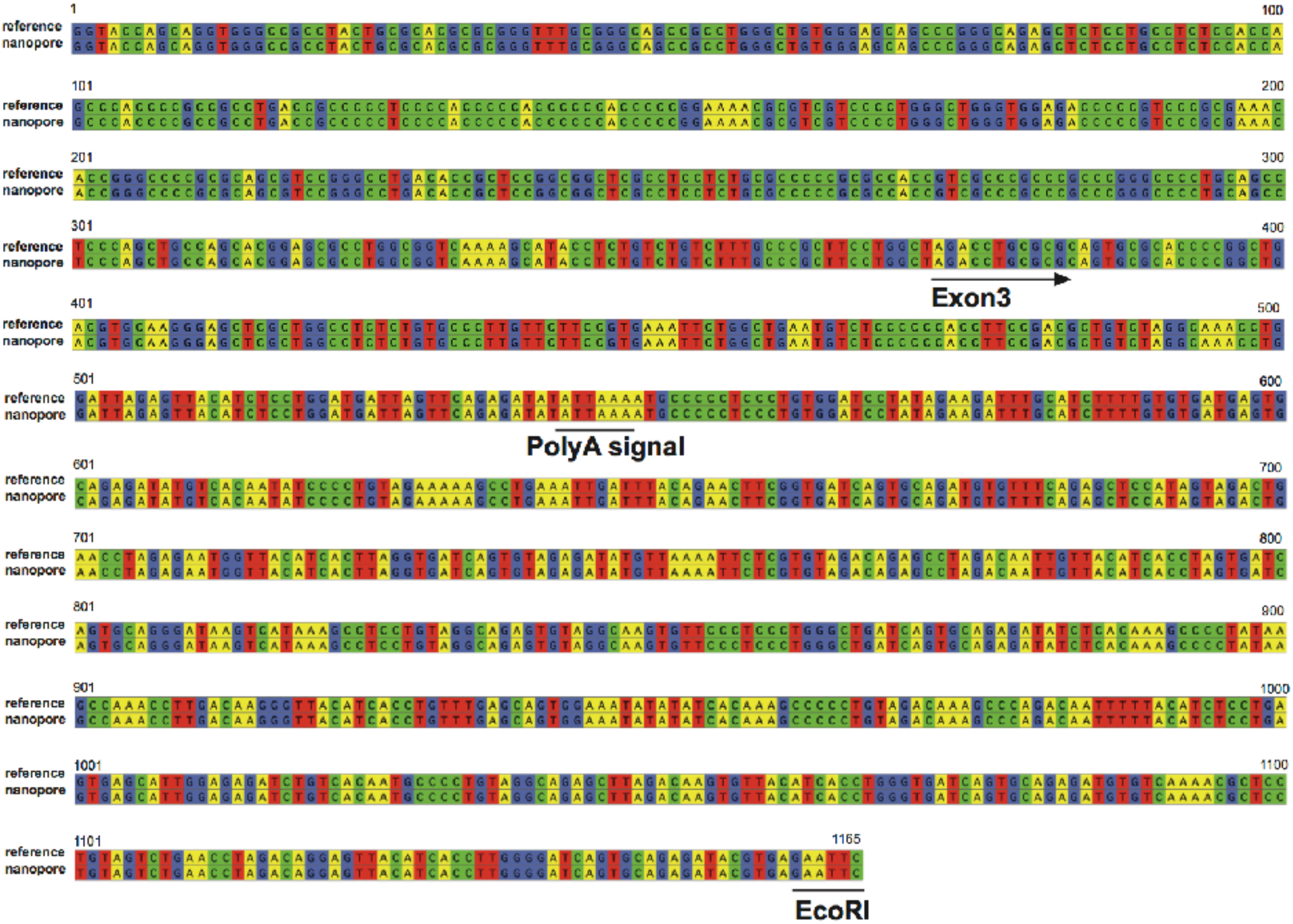


Number of reads	**128,171**
Total bases	**971,130,587**
Average read length	**7,577**

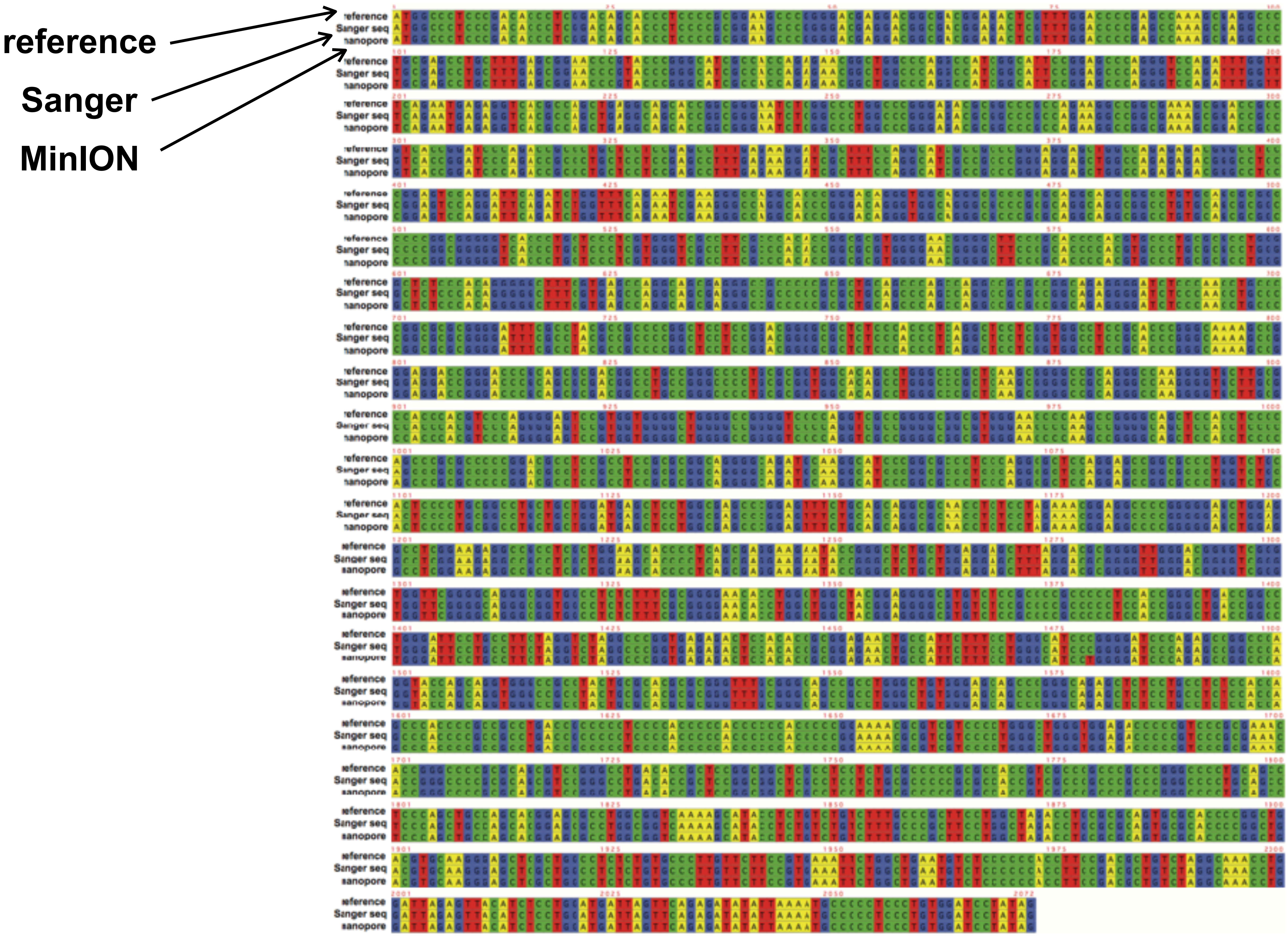
reference : CT476828.7

(Mitsuhashi et al. 2017 BioRxiv)

Sequence identity in pLAM region



DUX4 ORF was compared to Sanger-seq

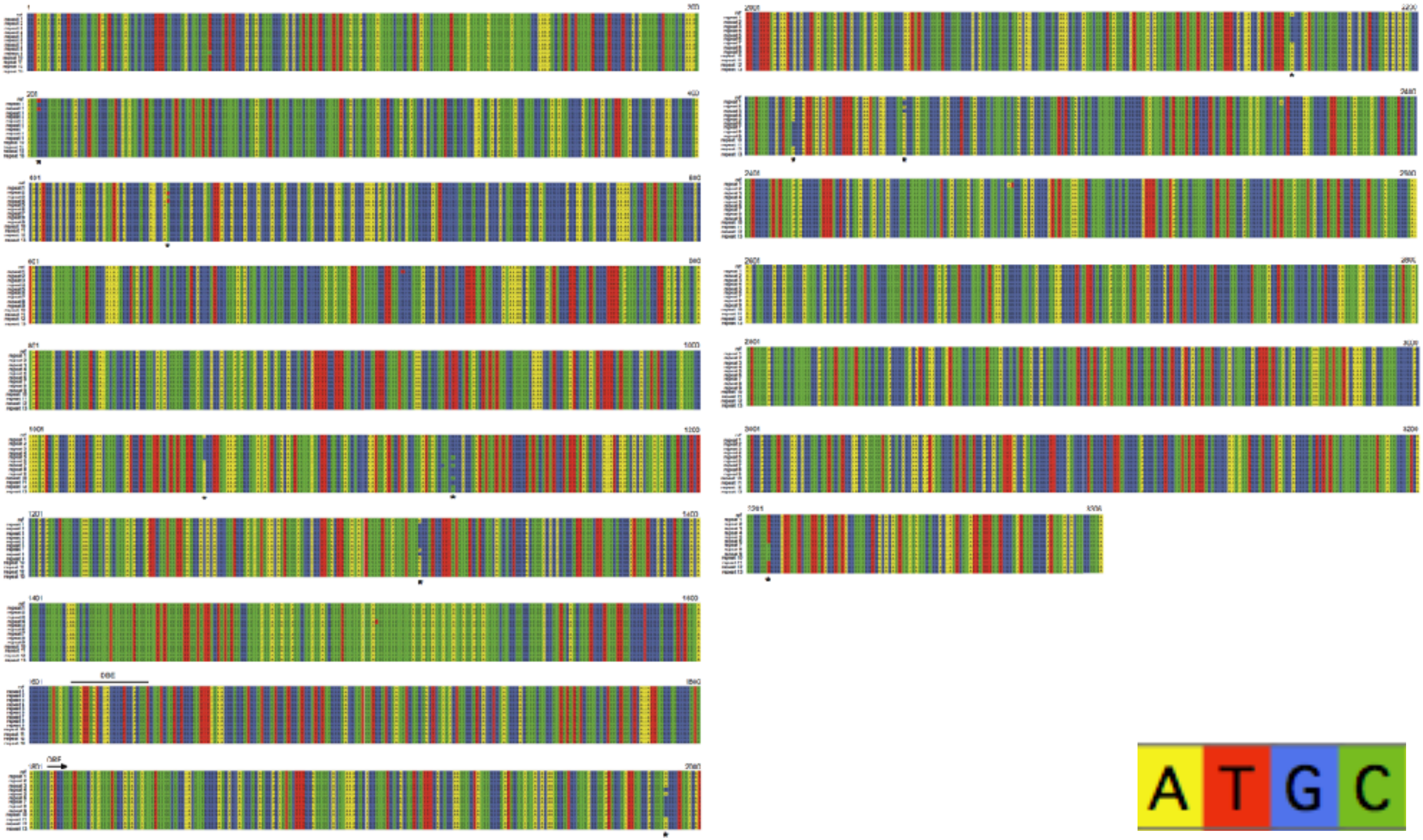


reference

Sanger

MinION

Sequence identity in 13 D4Z4 repeats



Consensus sequence was compared to the reference

Using the most common nucleotide

		MinION			
		A	C	G	T
ref	A		0	66	0
	C	2		2	29
	G	0	3		1
	T	0	20	12	

135 / 49,877(0.27%)

Genotyping using last-genotype

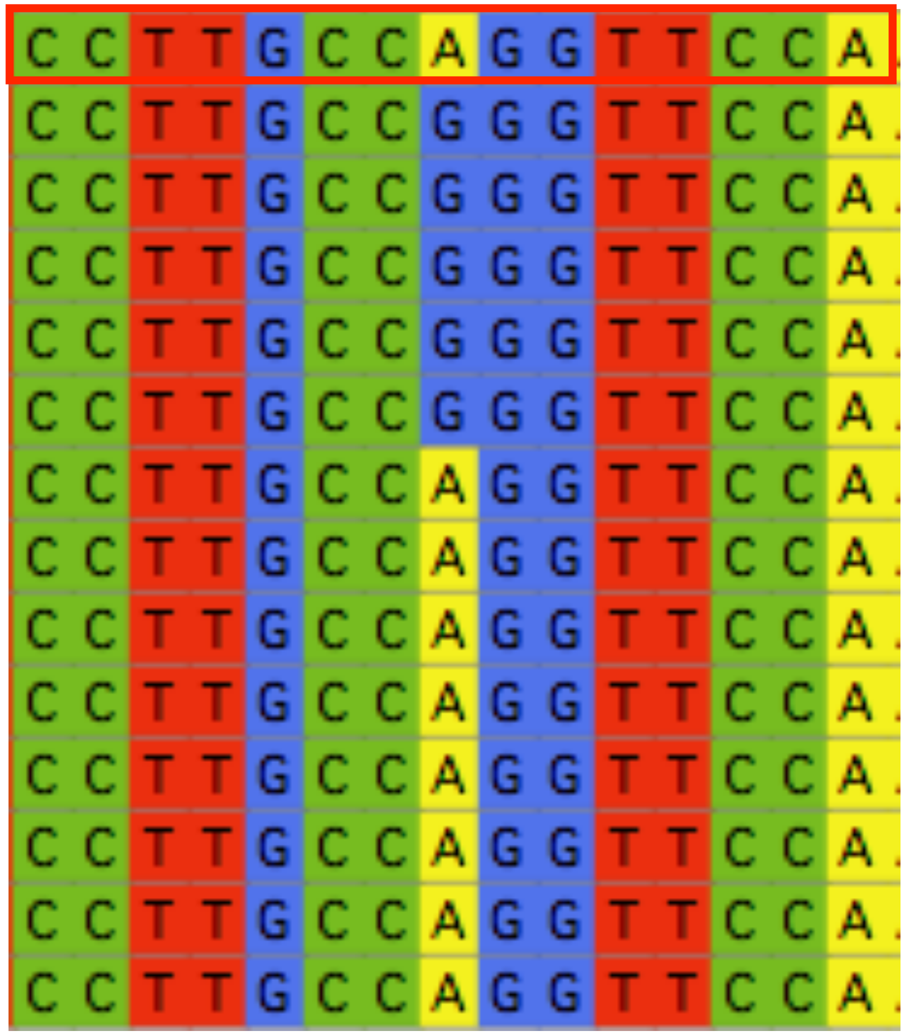
https://github.com/mcfrith/last-genotype

		MinION			
		A	C	G	T
ref	A		0	25	0
	C	0		13	13
	G	0	1		5
	T	2	15	1	

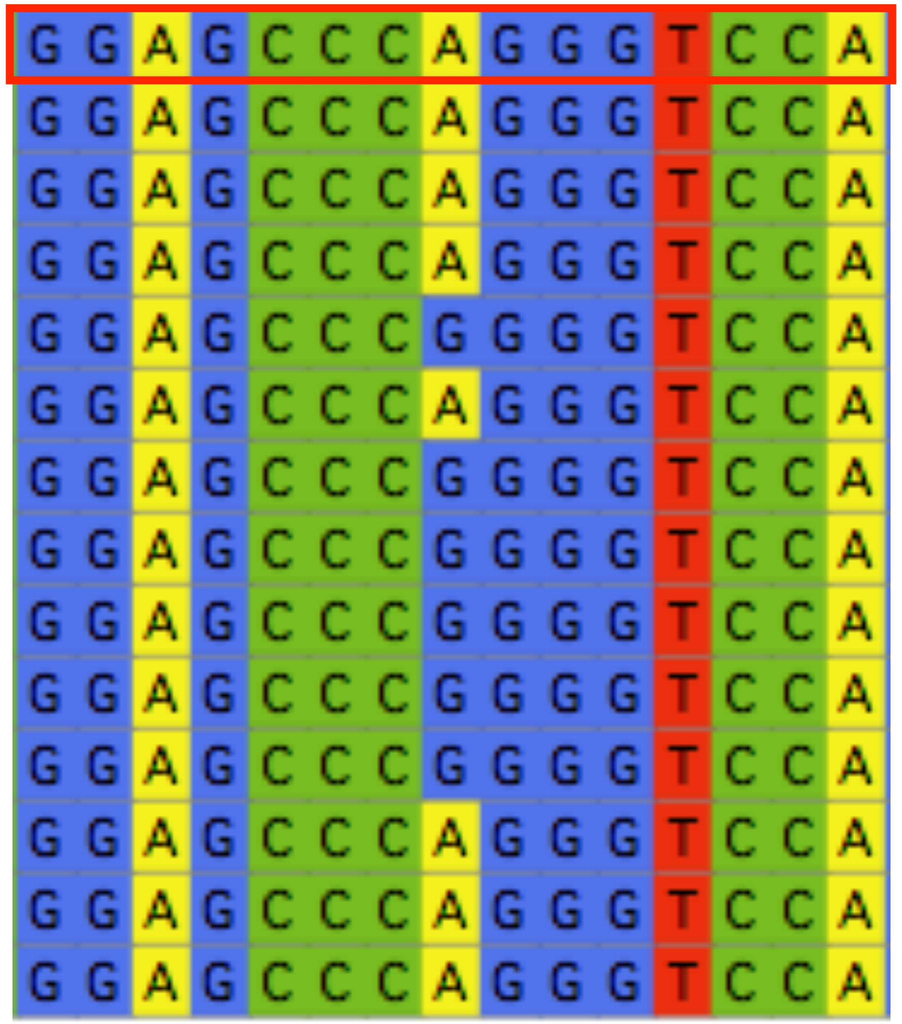
75 / 49,877(0.15%)

CCXGG are prone to be mis-base-called??

CCXGG



ref CCTTGCCAGGTTCCA
repeat1 CCTTGCCGGGTTCCA
CCTTGCCGGGTTCCA
CCTTGCCGGGTTCCA
CCTTGCCGGGTTCCA
CCTTGCCGGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
CCTTGCCAGGTTCCA
repeat13 CCTTGCCAGGTTCCA



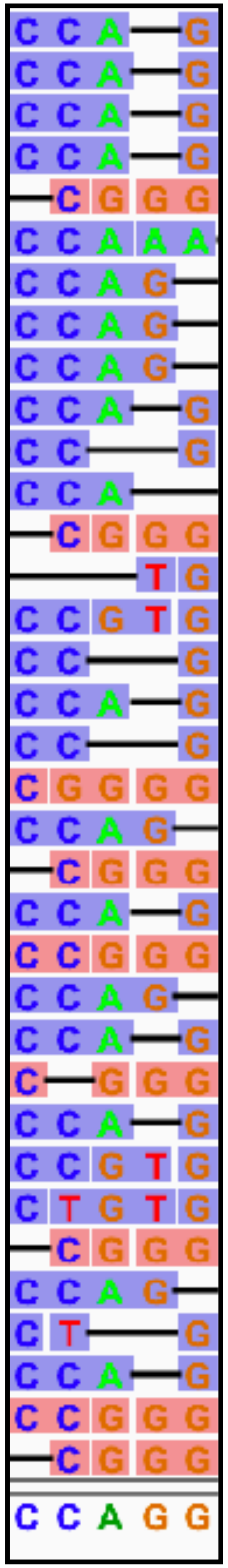
ref GGAGCCCAGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCGGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCAGGGTCCA
GGAGCCCAGGGTCCA

CCXGG errors:	most common nucleotide	117/135 (87%)
	last-genotype	49/62 (79%)

CCAGG > CCGGG : 32/49(65%)

last-genotype new version (as of Sep. 6. 2017)

https://github.com/mcfrith/last-genotype



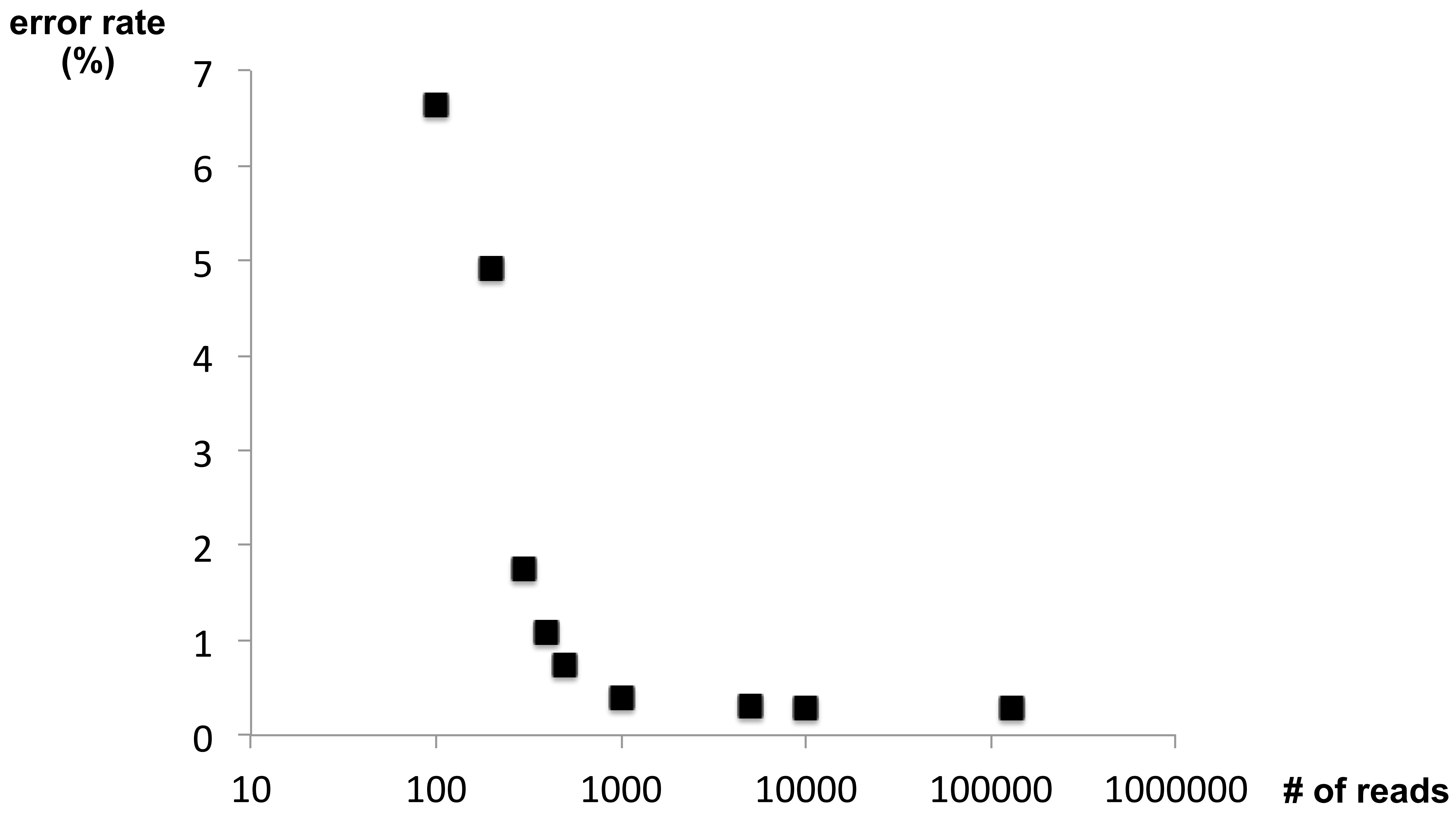
```
CCA-G
CCA-G
CCA-G
CCA-G
-CGGG
CCAAA
CCAG-
CCAG-
CCAG-
CCA-G
CC--G
CCA--
-CGGG
---TG
CCGTG
CC--G
CCA-G
CC--G
CGGGG
CCAG-
-CGGG
CCA-G
CCGGG
CCAG-
CCA-G
C-GGG
CCA-G
CCGTG
CTGTG
-CGGG
CCAG-
CT--G
CCA-G
CCGGG
-CGGG

CCAGG
```

		MinION			
		A	C	G	T
ref	A		0	1	0
	C	0		12	0
	G	0	1		5
	T	2	4	1	

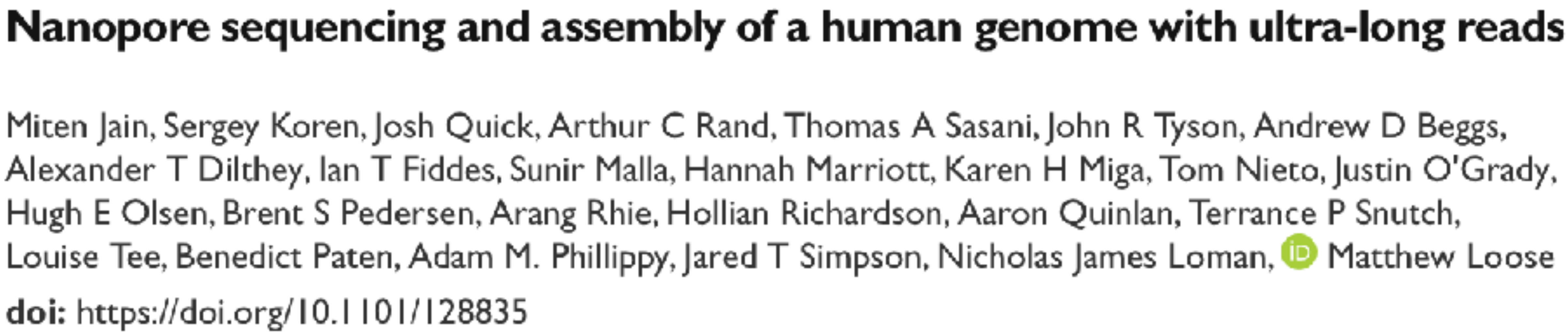
24 / 49,877(0.05%)

this slide was added in the last-minute

number of read and the consensus accuracy



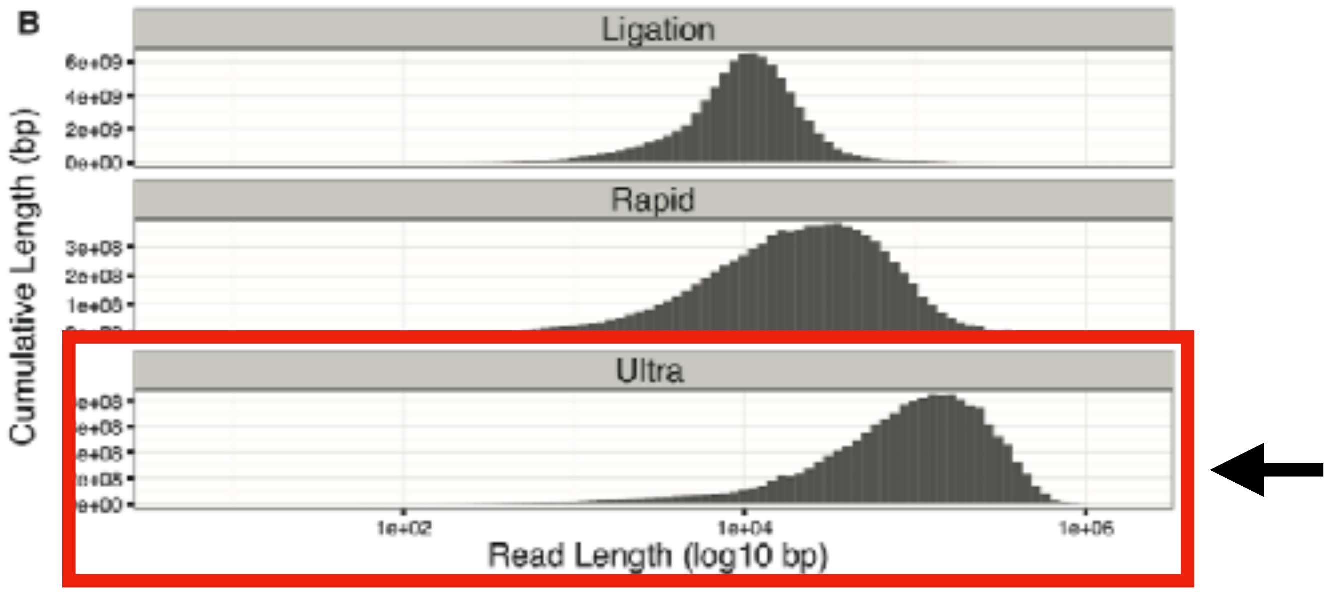
D4Z4 in whole human genome sequence dataset



Nanopore sequencing and assembly of a human genome with ultra-long reads

Miten Jain, Sergey Koren, Josh Quick, Arthur C Rand, Thomas A Sasani, John R Tyson, Andrew D Beggs, Alexander T Dilthey, Ian T Fiddes, Sunir Malla, Hannah Marriott, Karen H Miga, Tom Nieto, Justin O'Grady, Hugh E Olsen, Brent S Pedersen, Arang Rhie, Hollian Richardson, Aaron Quinlan, Terrance P Snutch, Louise Tee, Benedict Paten, Adam M. Phillippy, Jared T Simpson, Nicholas James Loman, Matthew Loose

doi: https://doi.org/10.1101/128835

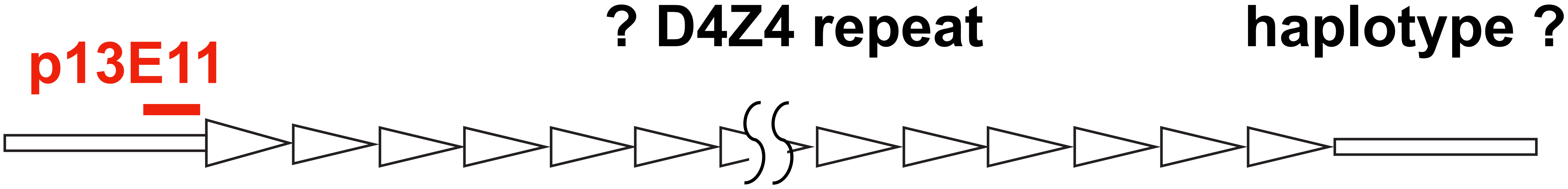


read #	**1,415,87**
data size	**23G**
average read length	**16,343**
maximum read length	**1,537,349**

X5 coverage of human genome

https://github.com/nanopore-wgs-consortium/NA12878

Obtaining D4Z4 repeat containing reads



p13E11

? D4Z4 repeat

haplotype ?

reads with p13E11

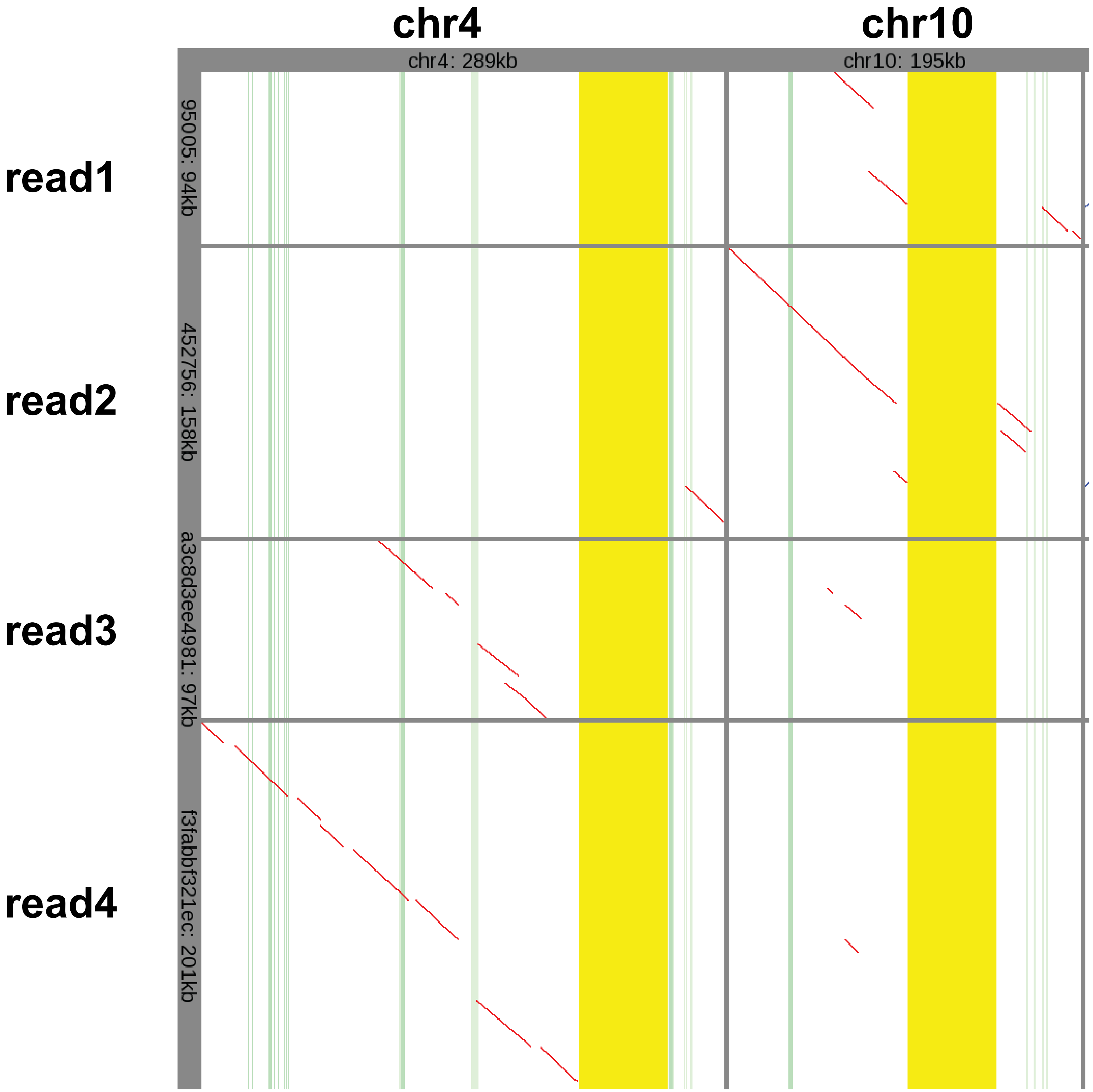
mapping to GRCh38

```
lastdb GRCh38
last-train
lastal | last-split
last-dotplot
```

mapping to single-D4Z4

```
lastdb single-D4Z4
last-train
lastal | last-split
last-dotplot
```

Reads from chr4 and chr10 can be distinguished



http://last.cbrc.jp/doc/last-dotplot.html

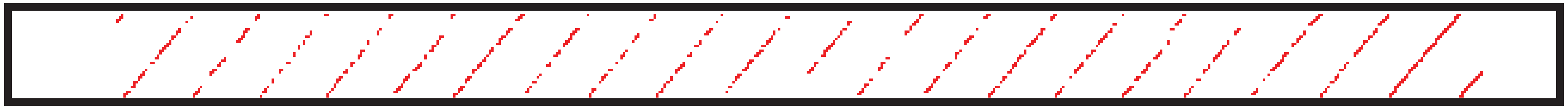
Number of the D4Z4 can be determined using last

Align to a single D4Z4 repeat using lastal

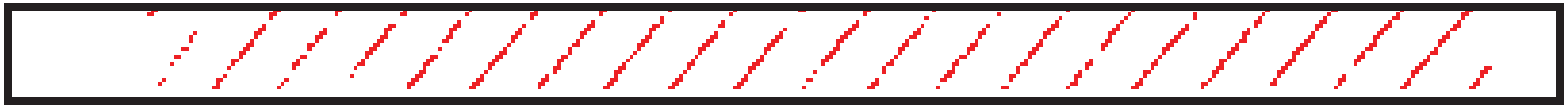
chr10

20 D4Z4 repeats

read1



read2

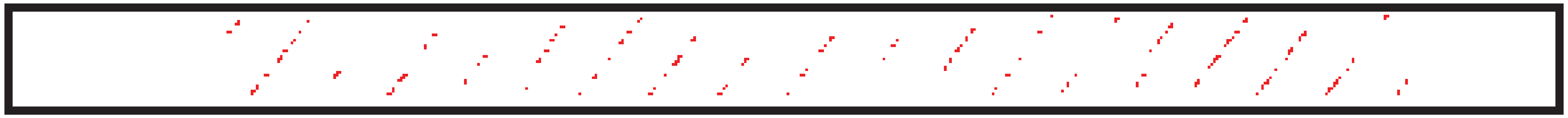


---

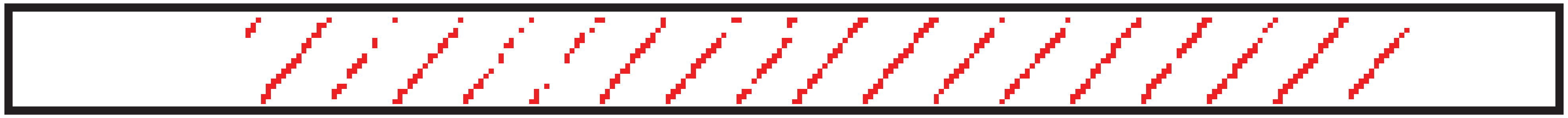
chr4

17 D4Z4 repeats

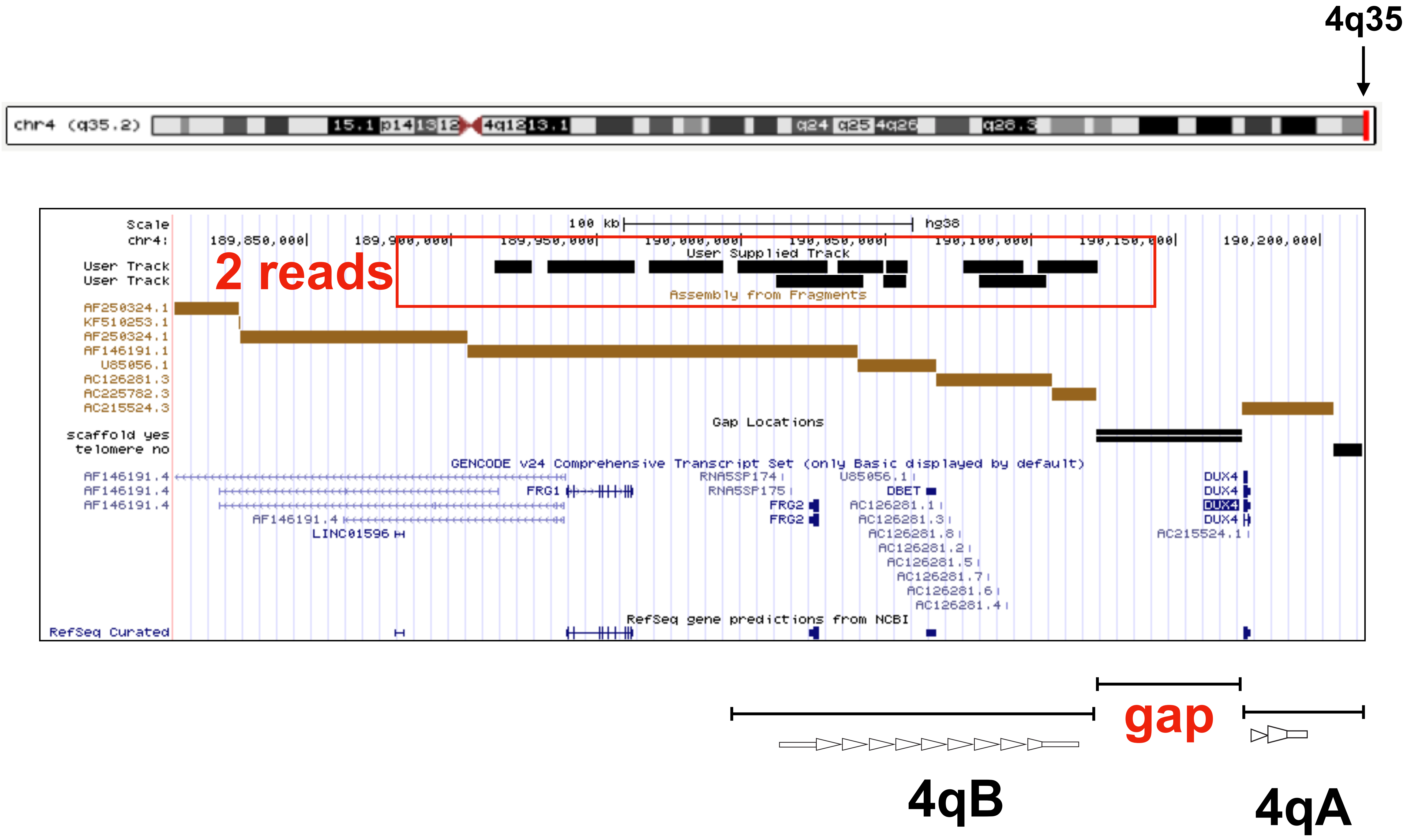
read3

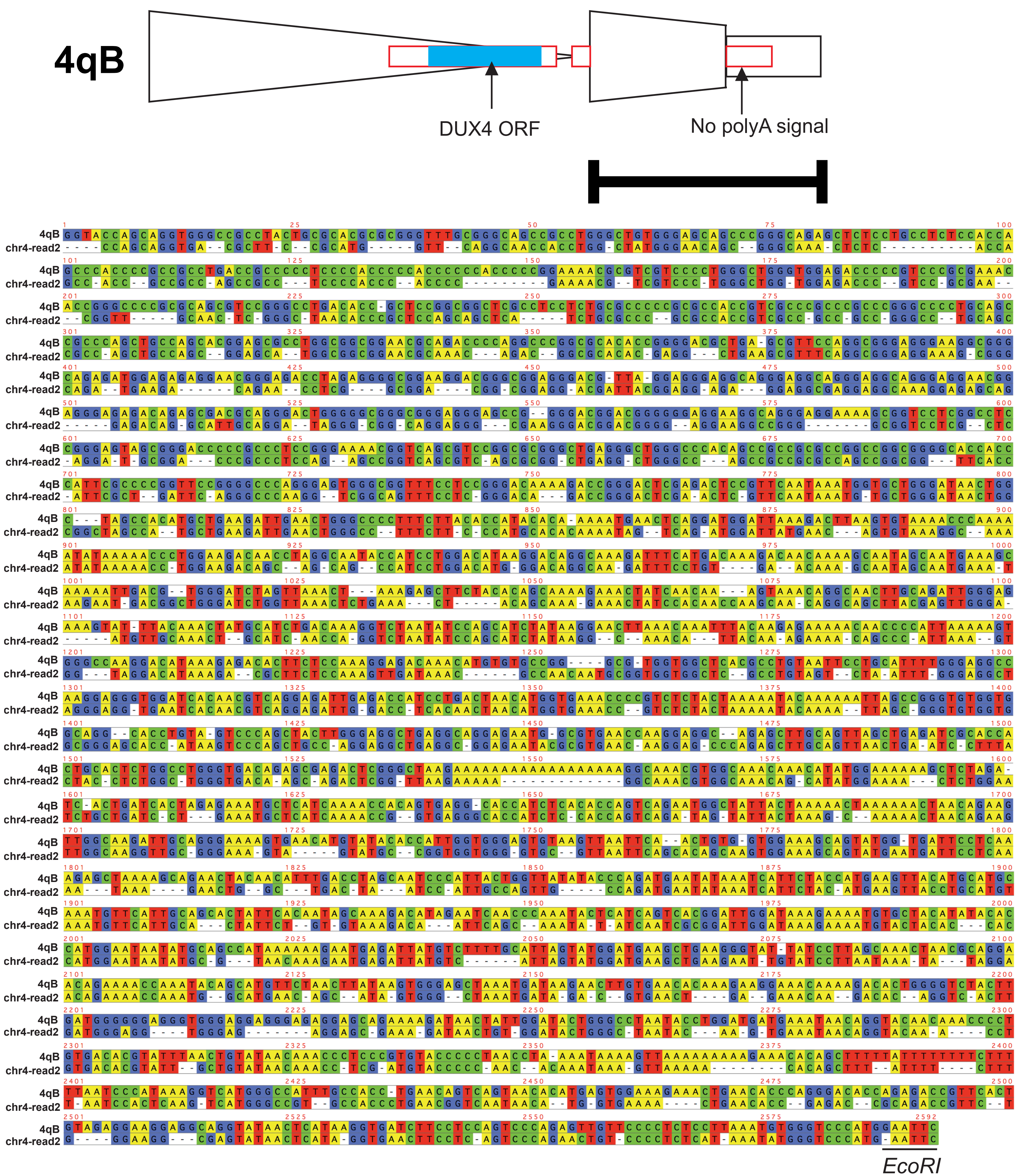


read4



chr4 D4Z4 on GRCh38 human reference genome




4qB
DUX4 ORF
No polyA signal
4qB
chr4-read2
EcoRI

Possible future diagnostic tool for FSHD??

1. **number of repeat** ✓
2. **distinguish homologous repeats** ✓
3. **haplotype** ✓

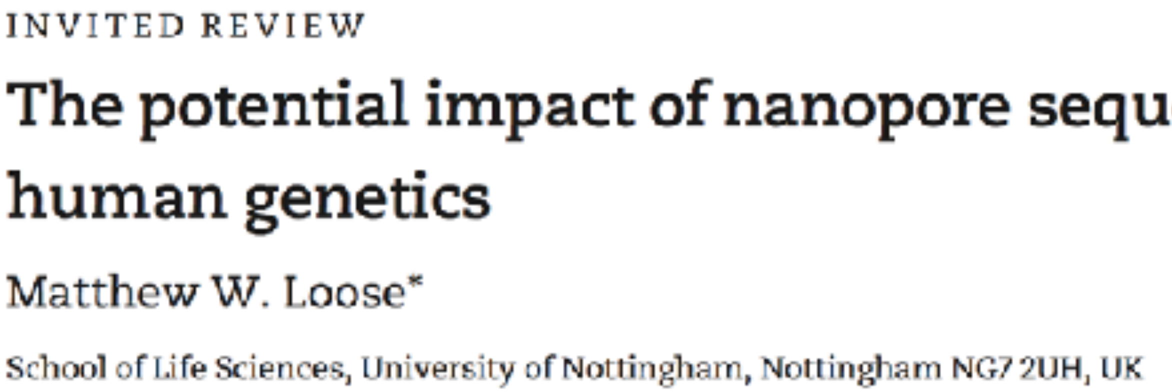
😊 → **Substitute for the Southernblotting?**

Why MinION?

There are many potential uses in the medical settings

Infectious disease
Mendelian disease
Autoimmune disease
Cancer
etc…

Nanopore sequencing will expand our knowledge on the pathomechanism of the human disease

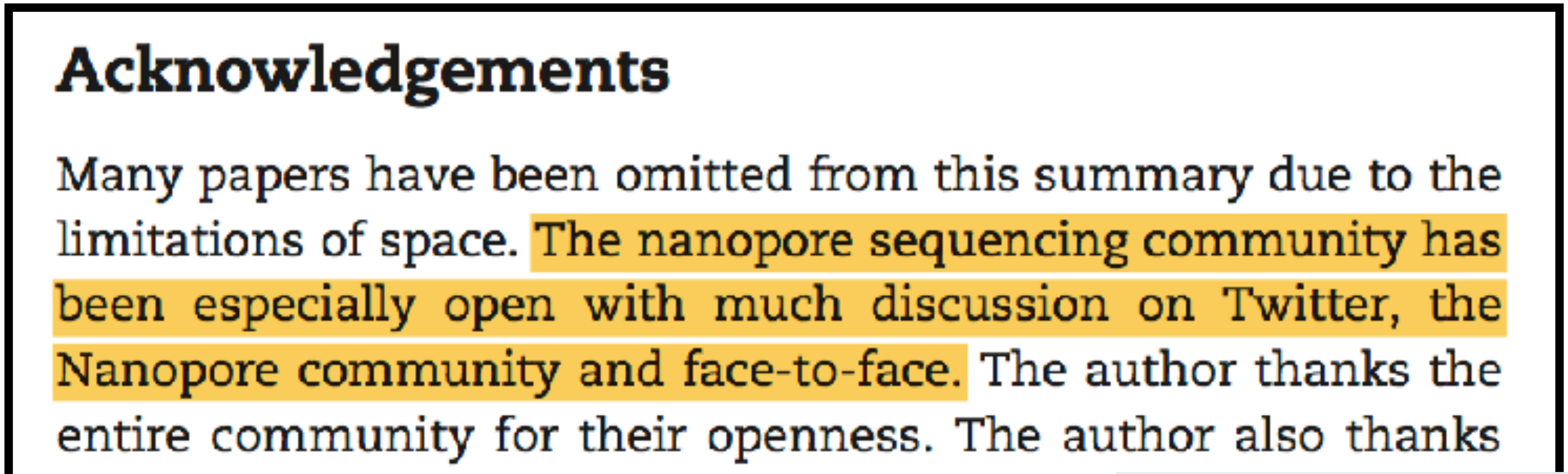


INVITED REVIEW

The potential impact of nanopore sequencing on human genetics

Matthew W. Loose*

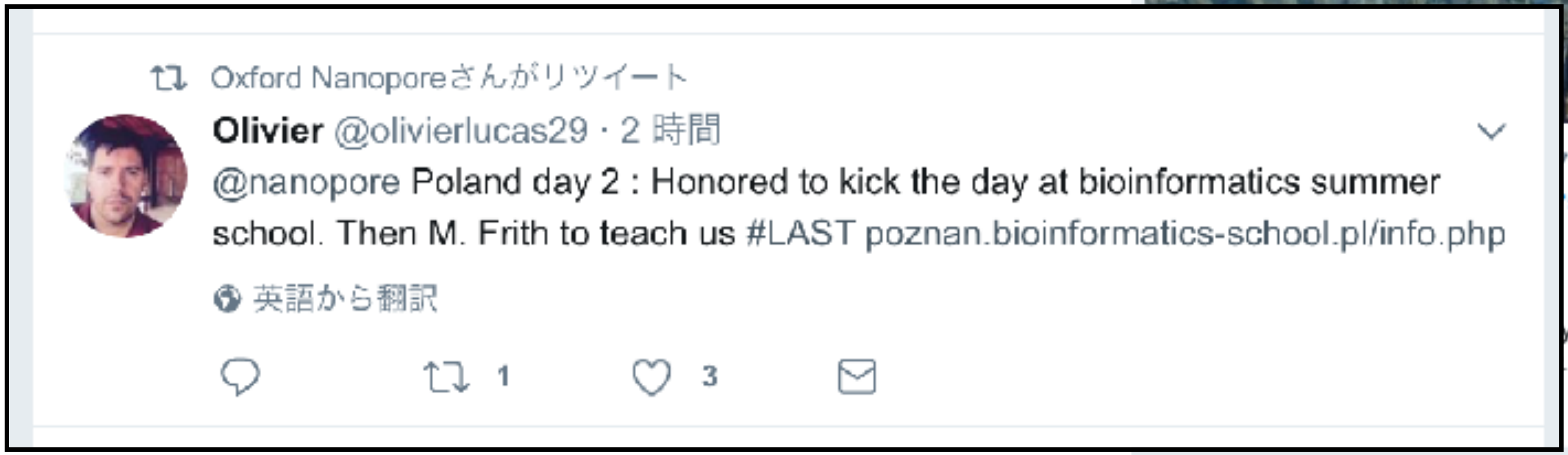
School of Life Sciences, University of Nottingham, Nottingham NG7 2UH, UK



Acknowledgements

Many papers have been omitted from this summary due to the limitations of space. The nanopore sequencing community has been especially open with much discussion on Twitter, the Nanopore community and face-to-face. The author thanks the entire community for their openness. The author also thanks Teri Evans for comments on this manuscript.

(Matthew W Loose. 2017. Hum Mol Genet)



Olivier @olivierlucas29 · 2 時間

@nanopore Poland day 2 : Honored to kick the day at bioinformatics summer school. Then M. Frith to teach us #LAST poznan.bioinformatics-school.pl/info.php

ysis: MinION, PromethION, GridION, VolTRAX.
. Metrichor Ltd for applied.

Acknowledgement

Tokai University

Hiroaki Mitsuhashi
So Nakagawa
Mahoko Ueda Takahashi
Tadashi Imanishi

AIST
University of Tokyo

Martin C Frith