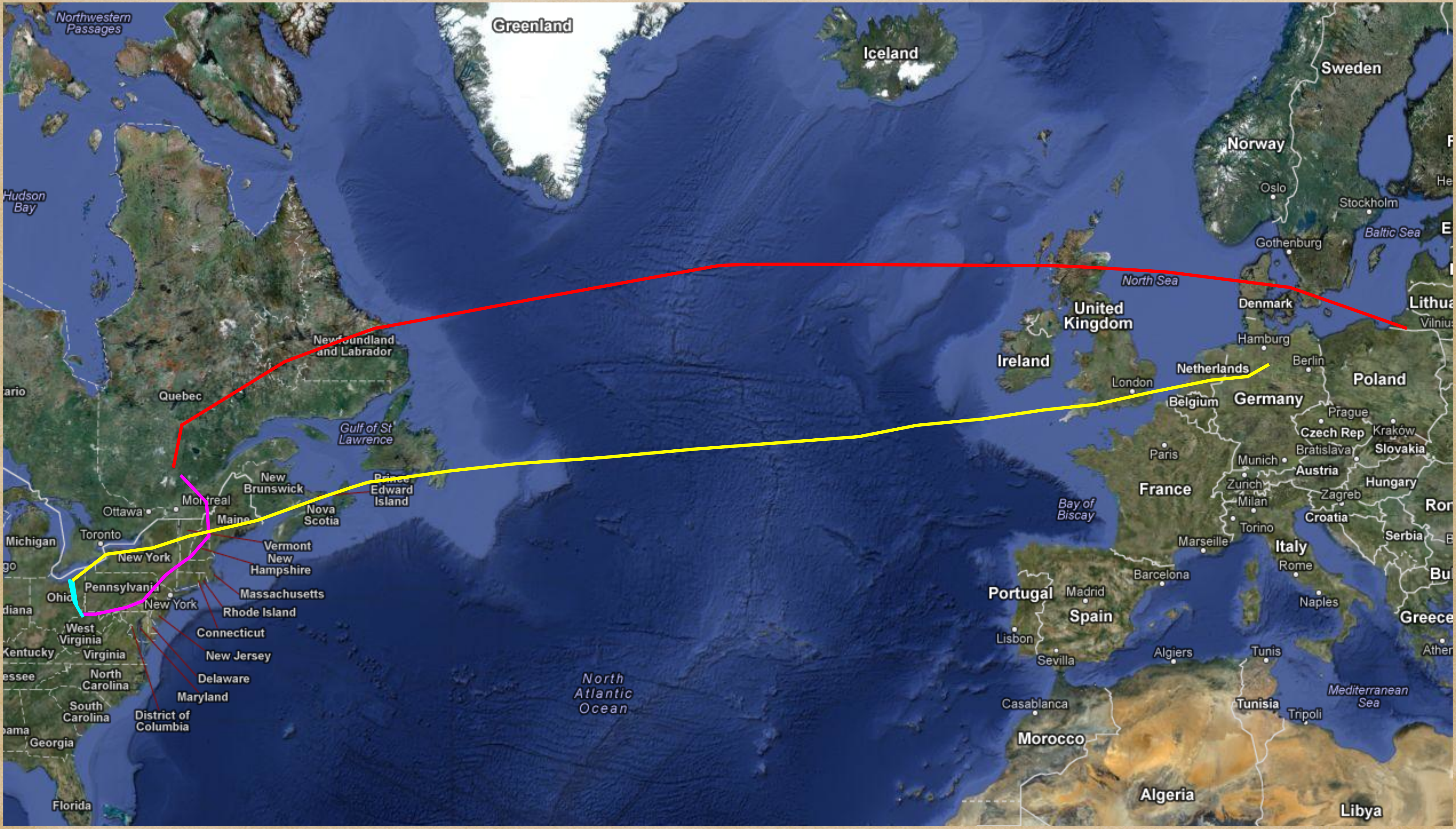
Nanopore Sequencing for Genotyping Pathogens of Tropical Diseases

Wojciech Makałowski
Institute of Bioinformatics, University of Muenster
and
Department of Computational Biology,
the University of Tokyo

http://bioinformatics.uni-muenster.de

Poznań -> Montreal ->Bethesda -> State College -> Münster -> Kashiwa

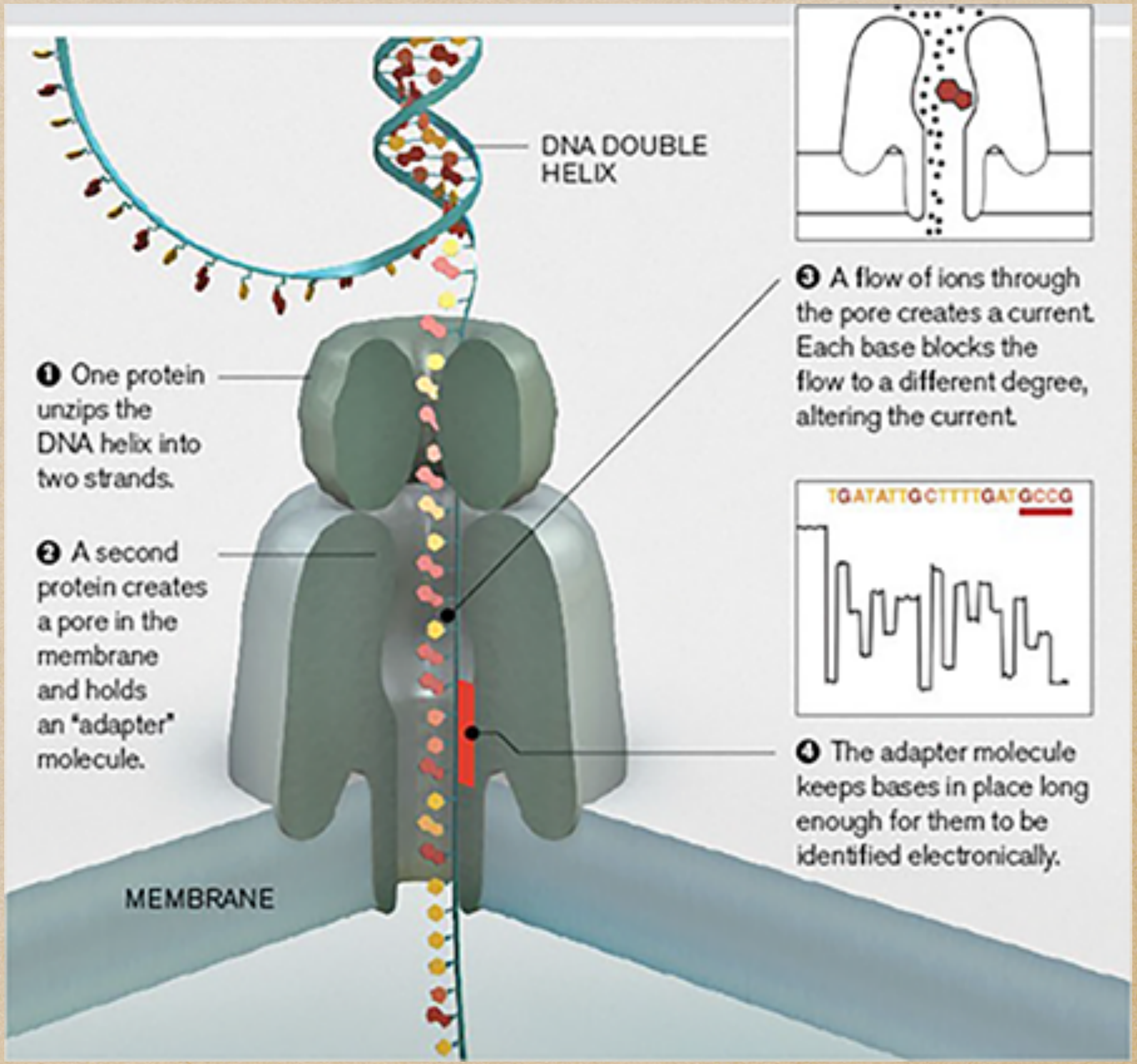


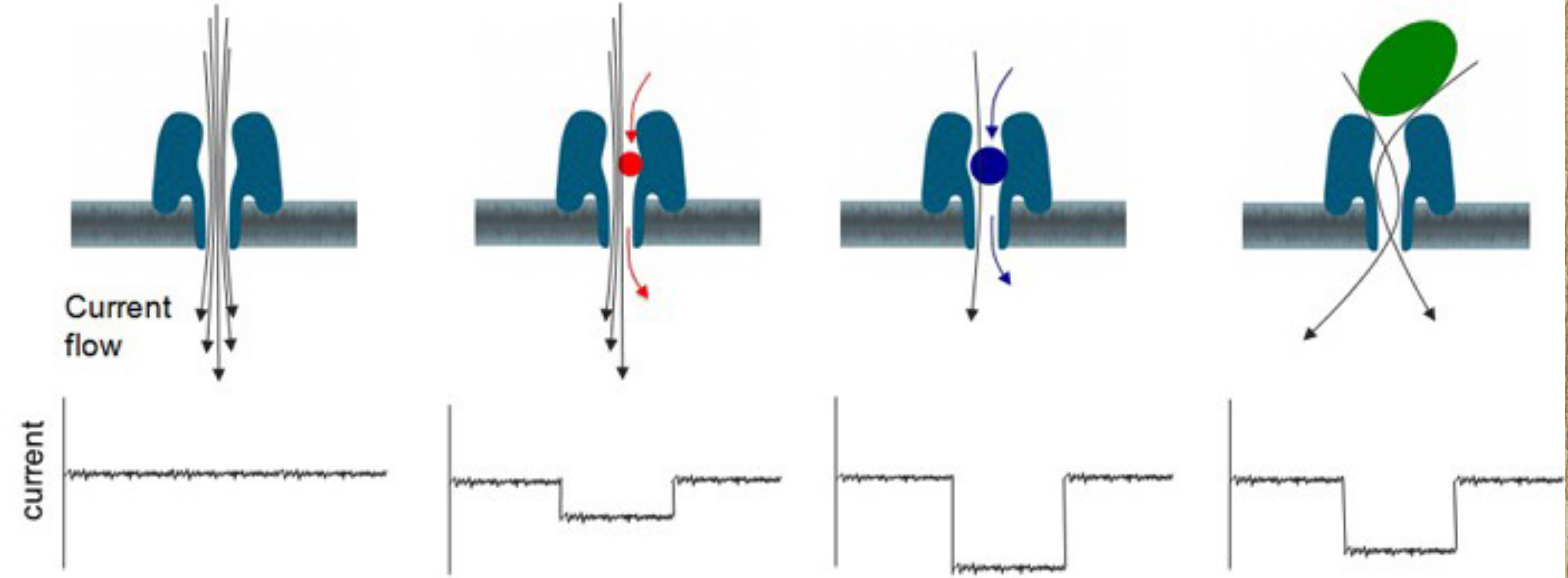
Third generation sequencing

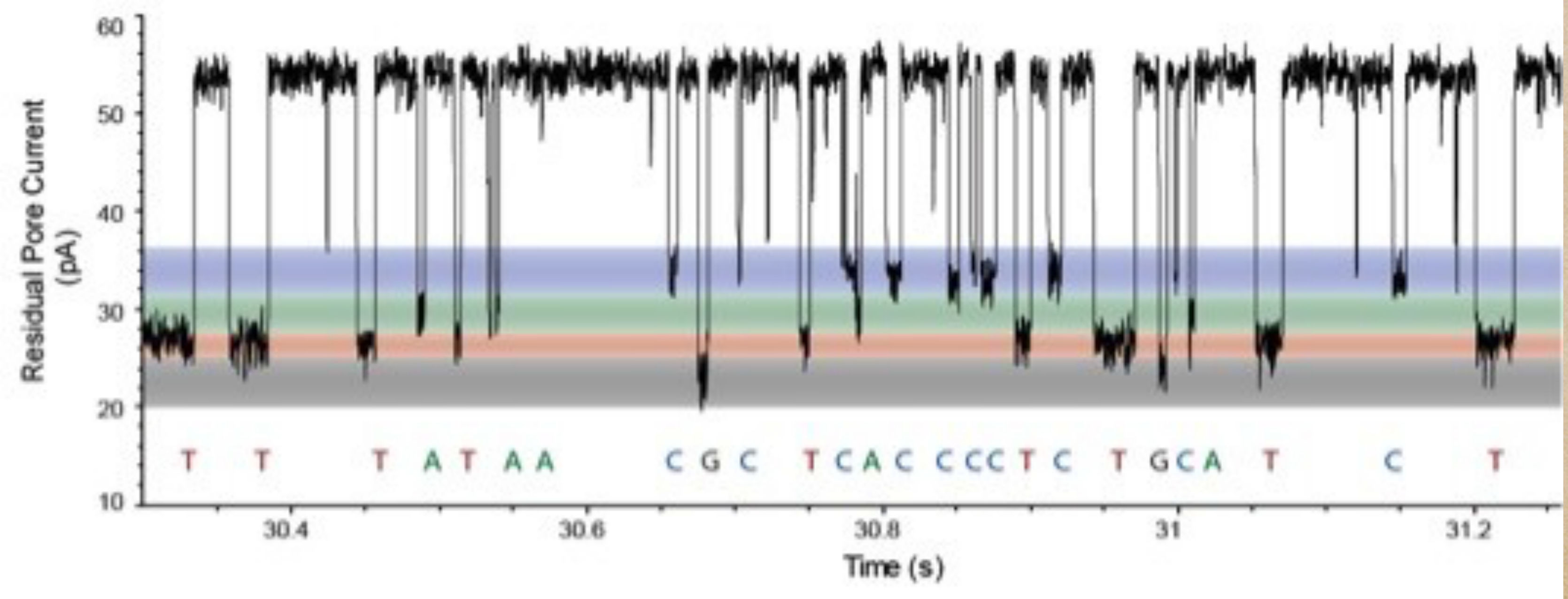
Single molecule sequencing: MinION by Oxford Nanopore

MinION basics

- Read length: > 10 Kb (Phage λ DNA, 50 Kb)
- # of pore : 512 (MinION at early access)
- Read speed: 8 bases to 20 bases/sec
- Run time: max 6 hours (MinION at early access)
- Error rate ≈ 20-30%
- 10,000 reads/run (?)

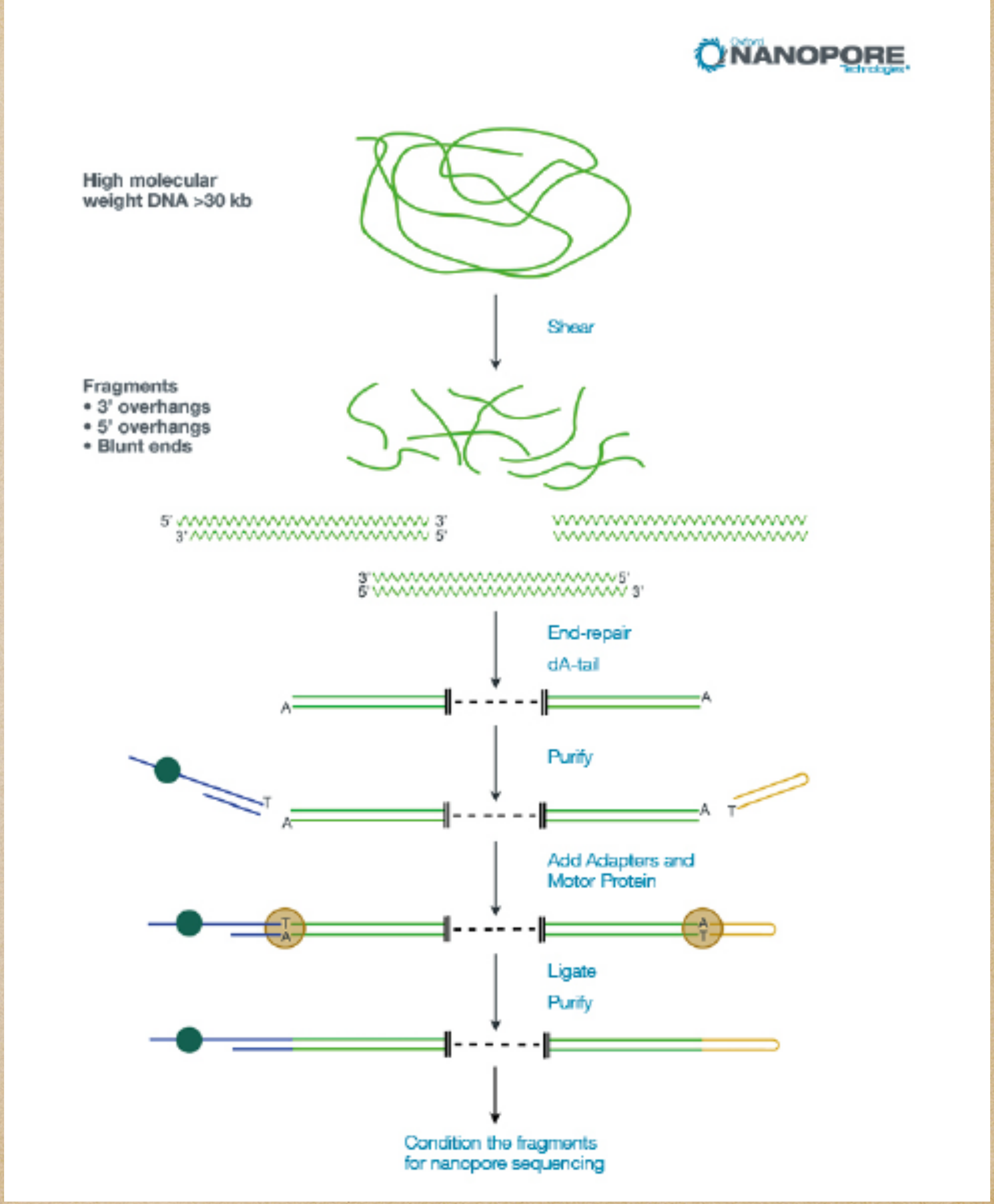



Current flow
current


Residual Pore Current (pA)
60
50
40
30
20
10
T T T A T A A C G C T C A C C C C T C T G C A T C T
30.4
30.6
30.8
31
31.2
Time (s)

MinION Access Program

Easy, standard template preparation



METRICHOR

Metrichor is an on-demand, cloud-based, bioinformatics data analysis platform. For the MAP, Metrichor supports Oxford Nanopore basecalling software for DNA sequencing applications. Basecalling may be made available locally to participants unable to use the cloud for regulatory reasons. The user will need to install a metrichor client on their computer.

MinKNOW

MinKNOW is the software that controls the MinION. It carries out several core data tasks and can be used to change experimental workflows or parameters. MinKNOW runs on the user's computer and is downloadable from a link provided.

MinION

Nanopore sensing is carried out on the sensor chip, contained in the flow cell inside the MinION device. Data is processed by an Application-Specific Integrated Circuit (ASIC) also in the flow cell and processed in real time by the MinKNOW software.

MinION dataflow

MinION

Nanopore sensing is carried out on the sensor chip, contained in the flow cell inside the MinION device. Data is processed by an Application-Specific Integrated Circuit (ASIC) also in the flow cell and processed in real time by the MinKNOW software

MinKNOW

MinKNOW is the software that controls the MinION. It carries out several core data tasks and can be used to change experimental workflows or parameters. MinKNOW runs on the user's computer.

METRICHOR

Metrichor is an on-demand, cloud-based, bioinformatics data analysis platform. It supports Oxford Nanopore base calling software. Base calling may be made available locally .

MiniKNOW - Data Render

Native Client - MinKNOW 0.45.1.6 b201406111512 - Running MAP_Sequencing_Run.py

Status ●
Acquisition ●
Analysis ●
Write ●

Exp. Time 21621s (stopped 39619s ago)
Yield 107402720
Sample ID BN_011_34887GT
Data Set nanopore-HP_BN_011_34887GT_2337_1

Asic Status ●
Asic 41.2°C
Heatsink 37.875°C
MinION ID MN02219

Bias Voltage -120mV
Asic ID 29679

Bias Voltage Offset 0
Bias Voltage Gain 5
Writing Raw Channels 0->1 (1 channels)

Channels in Reads: 212
Complete Reads: 53169
Events in Reads: 80761019

User Messages

```
16:18:52 Setting view to ChannelStateHistory
16:19:01 Selecting channels ..
16:21:51 ...finished selection.
16:21:56 run report deactivated
16:23:36 System ready
16:23:36 Setting view to GeneralInfo
16:23:52 Performing availability scan every 60
seconds
22:23:59 run report deactivated
```

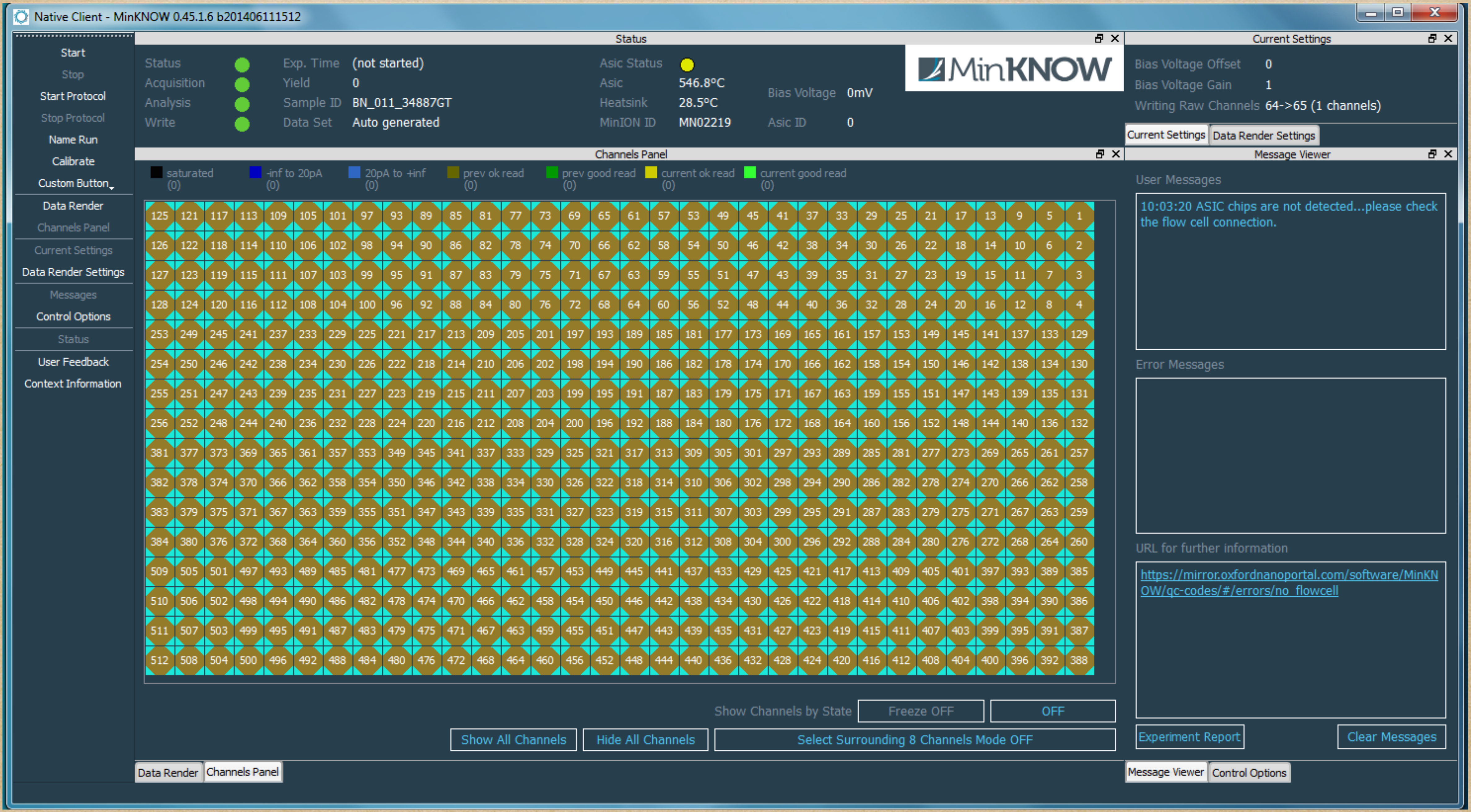
Error Messages

```
json_respond   EngineStatusError
EngineStatusError  u ERROR  ExecutionEngine
board_command_ex failed    board_command_ex
failed with ErrorFailedToExecute  Failed to execute
remote procedure      (2)
10:47:51 not calibrated
17:52:09 Buffer overflow - Data writer too slow
19:56:22 Heat sink temperature has not reached
37 degree Celsius
```

URL for further information

https://mirror.oxfordnanoportal.com/software/MinKNOW/qc-codes/#/errors/heatsink_temp_too_high

MiniKNOW - Channels Panel



Status ●
Acquisition ●
Analysis ●
Write ●

Exp. Time (not started)
Yield 0
Sample ID BN_011_34887GT
Data Set Auto generated

Asic Status ●
Asic 546.8°C
Heatsink 28.5°C
MinION ID MN02219

Bias Voltage 0mV
Asic ID 0

Bias Voltage Offset 0
Bias Voltage Gain 1
Writing Raw Channels 64->65 (1 channels)

■ saturated (0)
■ -inf to 20pA (0)
■ 20pA to +inf (0)
■ prev ok read (0)
■ prev good read (0)
■ current ok read (0)
■ current good read (0)

125	121	117	113	109	105	101	97	93	89	85	81	77	73	69	65	61	57	53	49	45	41	37	33	29	25	21	17	13	9	5	1
126	122	118	114	110	106	102	98	94	90	86	82	78	74	70	66	62	58	54	50	46	42	38	34	30	26	22	18	14	10	6	2
127	123	119	115	111	107	103	99	95	91	87	83	79	75	71	67	63	59	55	51	47	43	39	35	31	27	23	19	15	11	7	3
128	124	120	116	112	108	104	100	96	92	88	84	80	76	72	68	64	60	56	52	48	44	40	36	32	28	24	20	16	12	8	4
253	249	245	241	237	233	229	225	221	217	213	209	205	201	197	193	189	185	181	177	173	169	165	161	157	153	149	145	141	137	133	129
254	250	246	242	238	234	230	226	222	218	214	210	206	202	198	194	190	186	182	178	174	170	166	162	158	154	150	146	142	138	134	130
255	251	247	243	239	235	231	227	223	219	215	211	207	203	199	195	191	187	183	179	175	171	167	163	159	155	151	147	143	139	135	131
256	252	248	244	240	236	232	228	224	220	216	212	208	204	200	196	192	188	184	180	176	172	168	164	160	156	152	148	144	140	136	132
381	377	373	369	365	361	357	353	349	345	341	337	333	329	325	321	317	313	309	305	301	297	293	289	285	281	277	273	269	265	261	257
382	378	374	370	366	362	358	354	350	346	342	338	334	330	326	322	318	314	310	306	302	298	294	290	286	282	278	274	270	266	262	258
383	379	375	371	367	363	359	355	351	347	343	339	335	331	327	323	319	315	311	307	303	299	295	291	287	283	279	275	271	267	263	259
384	380	376	372	368	364	360	356	352	348	344	340	336	332	328	324	320	316	312	308	304	300	296	292	288	284	280	276	272	268	264	260
509	505	501	497	493	489	485	481	477	473	469	465	461	457	453	449	445	441	437	433	429	425	421	417	413	409	405	401	397	393	389	385
510	506	502	498	494	490	486	482	478	474	470	466	462	458	454	450	446	442	438	434	430	426	422	418	414	410	406	402	398	394	390	386
511	507	503	499	495	491	487	483	479	475	471	467	463	459	455	451	447	443	439	435	431	427	423	419	415	411	407	403	399	395	391	387
512	508	504	500	496	492	488	484	480	476	472	468	464	460	456	452	448	444	440	436	432	428	424	420	416	412	408	404	400	396	392	388

User Messages

10:03:20 ASIC chips are not detected...please check the flow cell connection.

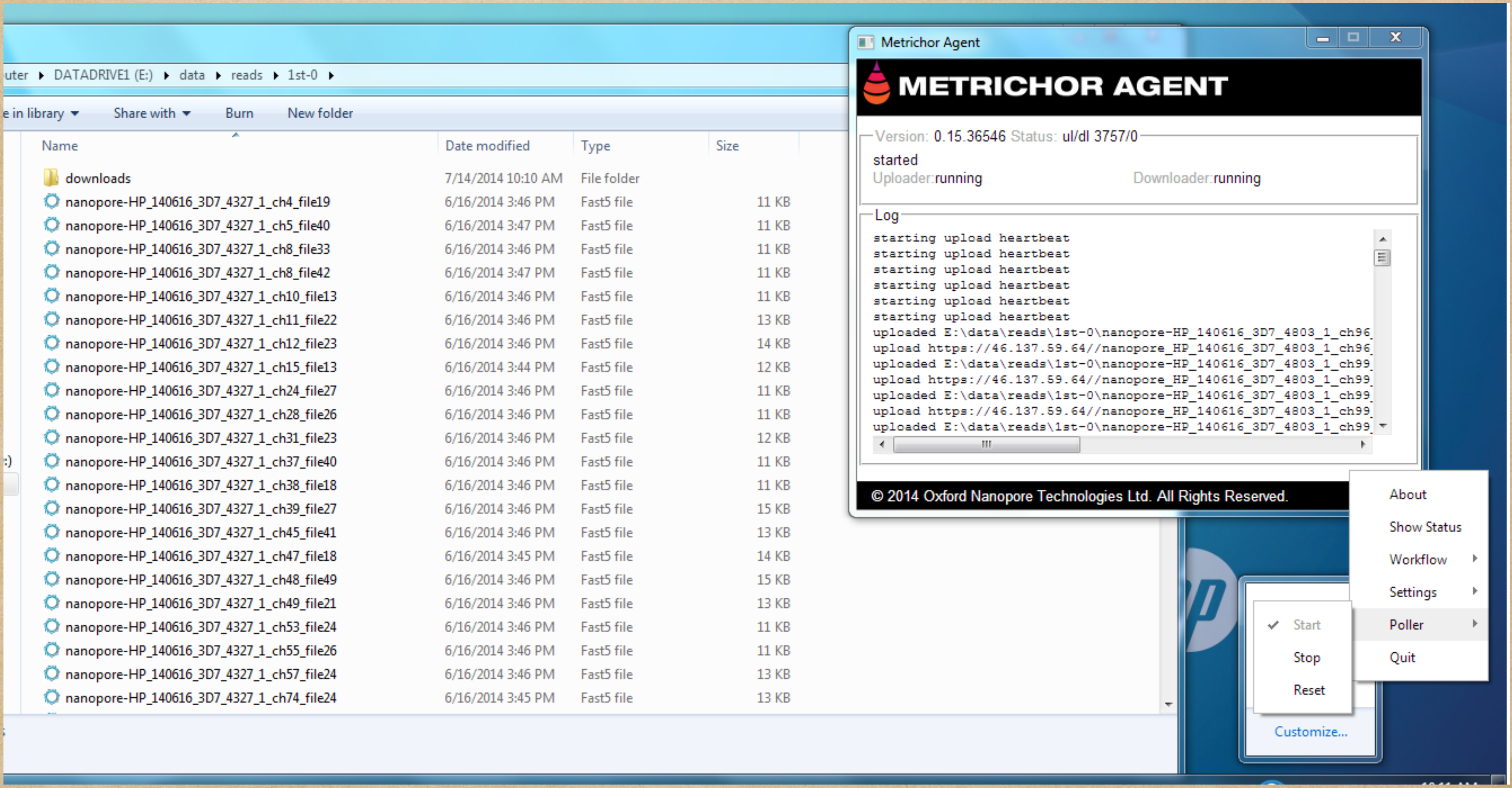
Error Messages

URL for further information

https://mirror.oxfordnanoportal.com/software/MinKNOW/qc-codes/#/errors/no_flowcell

Metrichor

Raw data, BaseCall data -> .Fast5 file



Name	Date modified	Type	Size
downloads	7/14/2014 10:10 AM	File folder	
nanopore-HP_140616_3D7_4327_1_ch4_file19	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch5_file40	6/16/2014 3:47 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch8_file33	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch8_file42	6/16/2014 3:47 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch10_file13	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch11_file22	6/16/2014 3:46 PM	Fast5 file	13 KB
nanopore-HP_140616_3D7_4327_1_ch12_file23	6/16/2014 3:46 PM	Fast5 file	14 KB
nanopore-HP_140616_3D7_4327_1_ch15_file13	6/16/2014 3:44 PM	Fast5 file	12 KB
nanopore-HP_140616_3D7_4327_1_ch24_file27	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch28_file26	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch31_file23	6/16/2014 3:46 PM	Fast5 file	12 KB
nanopore-HP_140616_3D7_4327_1_ch37_file40	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch38_file18	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch39_file27	6/16/2014 3:46 PM	Fast5 file	15 KB
nanopore-HP_140616_3D7_4327_1_ch45_file41	6/16/2014 3:46 PM	Fast5 file	13 KB
nanopore-HP_140616_3D7_4327_1_ch47_file18	6/16/2014 3:45 PM	Fast5 file	14 KB
nanopore-HP_140616_3D7_4327_1_ch48_file49	6/16/2014 3:46 PM	Fast5 file	15 KB
nanopore-HP_140616_3D7_4327_1_ch49_file21	6/16/2014 3:46 PM	Fast5 file	13 KB
nanopore-HP_140616_3D7_4327_1_ch53_file24	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch55_file26	6/16/2014 3:46 PM	Fast5 file	11 KB
nanopore-HP_140616_3D7_4327_1_ch57_file24	6/16/2014 3:46 PM	Fast5 file	13 KB
nanopore-HP_140616_3D7_4327_1_ch74_file24	6/16/2014 3:45 PM	Fast5 file	13 KB

Version: 0.15.36546 Status: ul/dl 3757/0

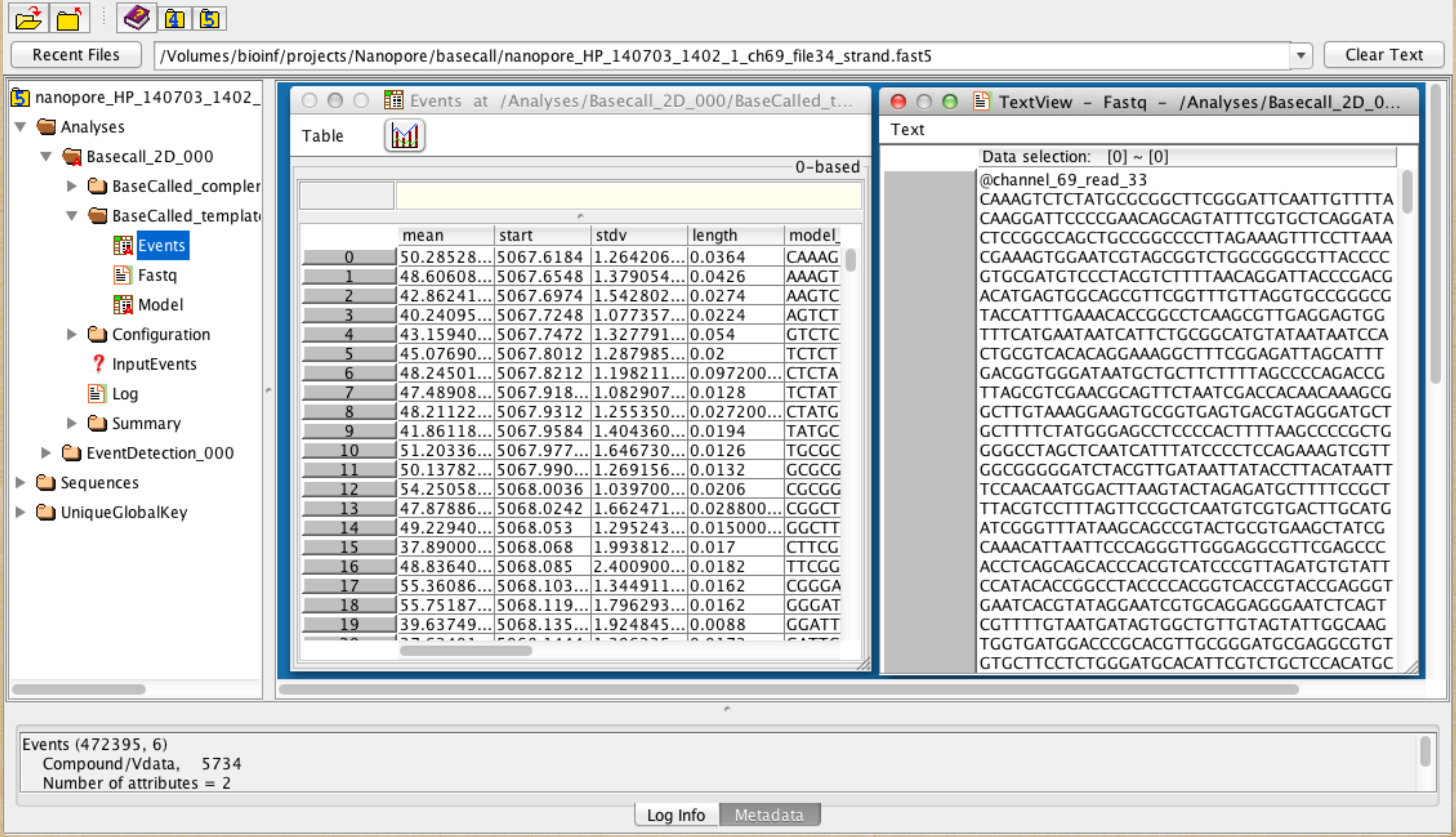
started

Uploader:running Downloader:running

Log

```
starting upload heartbeat
starting upload heartbeat
starting upload heartbeat
starting upload heartbeat
starting upload heartbeat
starting upload heartbeat
uploaded E:\data\reads\1st-0\nanopore-HP_140616_3D7_4803_1_ch96_
upload https://46.137.59.64//nanopore_HP_140616_3D7_4803_1_ch96_
uploaded E:\data\reads\1st-0\nanopore-HP_140616_3D7_4803_1_ch99_
upload https://46.137.59.64//nanopore_HP_140616_3D7_4803_1_ch99_
uploaded E:\data\reads\1st-0\nanopore-HP_140616_3D7_4803_1_ch99_
upload https://46.137.59.64//nanopore_HP_140616_3D7_4803_1_ch99_
uploaded E:\data\reads\1st-0\nanopore-HP_140616_3D7_4803_1_ch99_
```

HDFView



/Volumes/bioinf/projects/Nanopore/basecall/nanopore_HP_140703_1402_1_ch69_file34_strand.fast5

	mean	start	stdv	length	model_
0	50.28528...	5067.6184	1.264206...	0.0364	CAAAG
1	48.60608...	5067.6548	1.379054...	0.0426	AAAGT
2	42.86241...	5067.6974	1.542802...	0.0274	AAGTC
3	40.24095...	5067.7248	1.077357...	0.0224	AGTCT
4	43.15940...	5067.7472	1.327791...	0.054	GTCTC
5	45.07690...	5067.8012	1.287985...	0.02	TCTCT
6	48.24501...	5067.8212	1.198211...	0.097200...	CTCTA
7	47.48908...	5067.918...	1.082907...	0.0128	TCTAT
8	48.21122...	5067.9312	1.255350...	0.027200...	CTATG
9	41.86118...	5067.9584	1.404360...	0.0194	TATGC
10	51.20336...	5067.977...	1.646730...	0.0126	TGCGC
11	50.13782...	5067.990...	1.269156...	0.0132	GCGCG
12	54.25058...	5068.0036	1.039700...	0.0206	CGCGG
13	47.87886...	5068.0242	1.662471...	0.028800...	CGGCT
14	49.22940...	5068.053	1.295243...	0.015000...	GGCTT
15	37.89000...	5068.068	1.993812...	0.017	CTTCG
16	48.83640...	5068.085	2.400900...	0.0182	TTCGG
17	55.36086...	5068.103...	1.344911...	0.0162	CGGGA
18	55.75187...	5068.119...	1.796293...	0.0162	GGGAT
19	39.63749...	5068.135...	1.924845...	0.0088	GGATT

Data selection: [0] ~ [0]

```
@channel_69_read_33
CAAAGTCTCTATGCGCGGCTTCGGGATTCAATTGTTTTA
CAAGGATTCCCCGAACAGCAGTATTTCGTGCTCAGGATA
CTCCGGCCAGCTGCCGGCCCCTTAGAAAGTTTCCTTAAA
CGAAAGTGGAATCGTAGCGGTCTGGCGGGCGTTACCCC
GTGCGATGTCCCTACGTCTTTTAACAGGATTACCCGACG
ACATGAGTGGCAGCGTTCGGTTTGTTAGGTGCCGGGCG
TACCATTTGAAACACCGGCCTCAAGCGTTGAGGAGTGG
TTTCATGAATAATCATTCTGCGGCATGTATAATAATCCA
CTGCGTCACACAGGAAAGGCTTTCGGAGATTAGCATTT
GACGGTGGGATAATGCTGCTTCTTTTAGCCCCAGACCG
TTAGCGTCGAACGCAGTTCTAATCGACCACAACAAAGCG
GCTTGTAAAGGAAGTGCGGTGAGTGACGTAGGGATGCT
GCTTTTCTATGGGAGCCTCCCCACTTTTAAGCCCCGCTG
GGGCCTAGCTCAATCATTTATCCCCTCCAGAAAGTCGTT
GGCGGGGGATCTACGTTGATAATTATACCTTACATAATT
TCCAACAATGGACTTAAGTACTAGAGATGCTTTTCCGCT
TTACGTCCTTTAGTTCCGCTCAATGTCGTGACTTGCATG
ATCGGGTTTATAAGCAGCCGTACTGCGTGAAGCTATCG
CAAACATTAATTCCCAGGGTTGGGAGGCGTTCGAGCCC
ACCTCAGCAGCACCCACGTCATCCCGTTAGATGTGTATT
CCATACACCGGCCTACCCCACGGTCACCGTACCGAGGGT
GAATCACGTATAGGAATCGTGCAGGAGGGAATCTCAGT
CGTTTTGTAATGATAGTGGCTGTTGTAGTATTGGCAAG
TGGTGATGGACCCGCACGTTGCGGGATGCGAGGCGTGT
GTGCTTCCTCTGGGATGCACATTCGTCTGCTCCACATGC
```

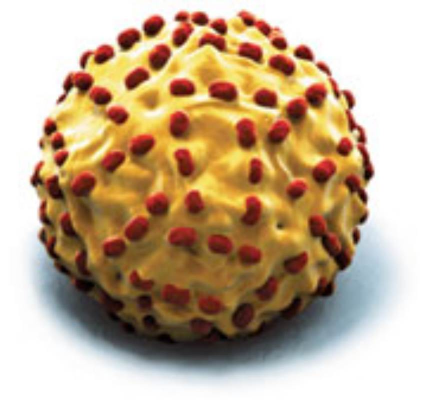
```
Events (472395, 6)
  Compound/Vdata,    5734
  Number of attributes = 2
```

Advantage of nanopore technology

- Label-free
- Single molecule, long reads analysis
- Disposable; autoclavable after the use
- Portable; requires no pre-installation of any instruments

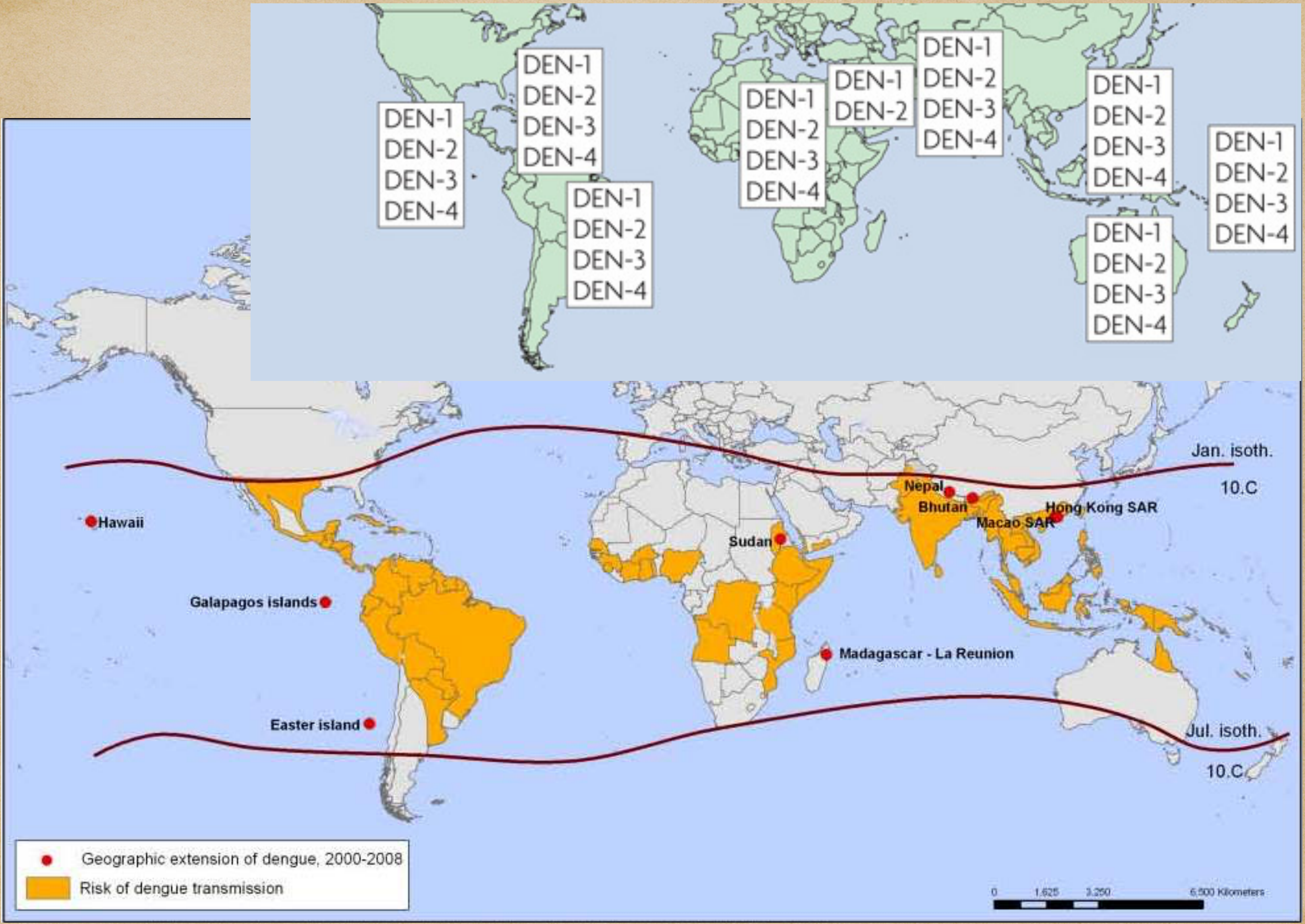
Potential for tropical diseases research and diagnostics

- In many countries where tropical diseases prevail
  - no conventional sequencer/ PCR instruments are available
  - shortage of well-trained technical staff
  - Needs for handling potentially dangerous pathogens

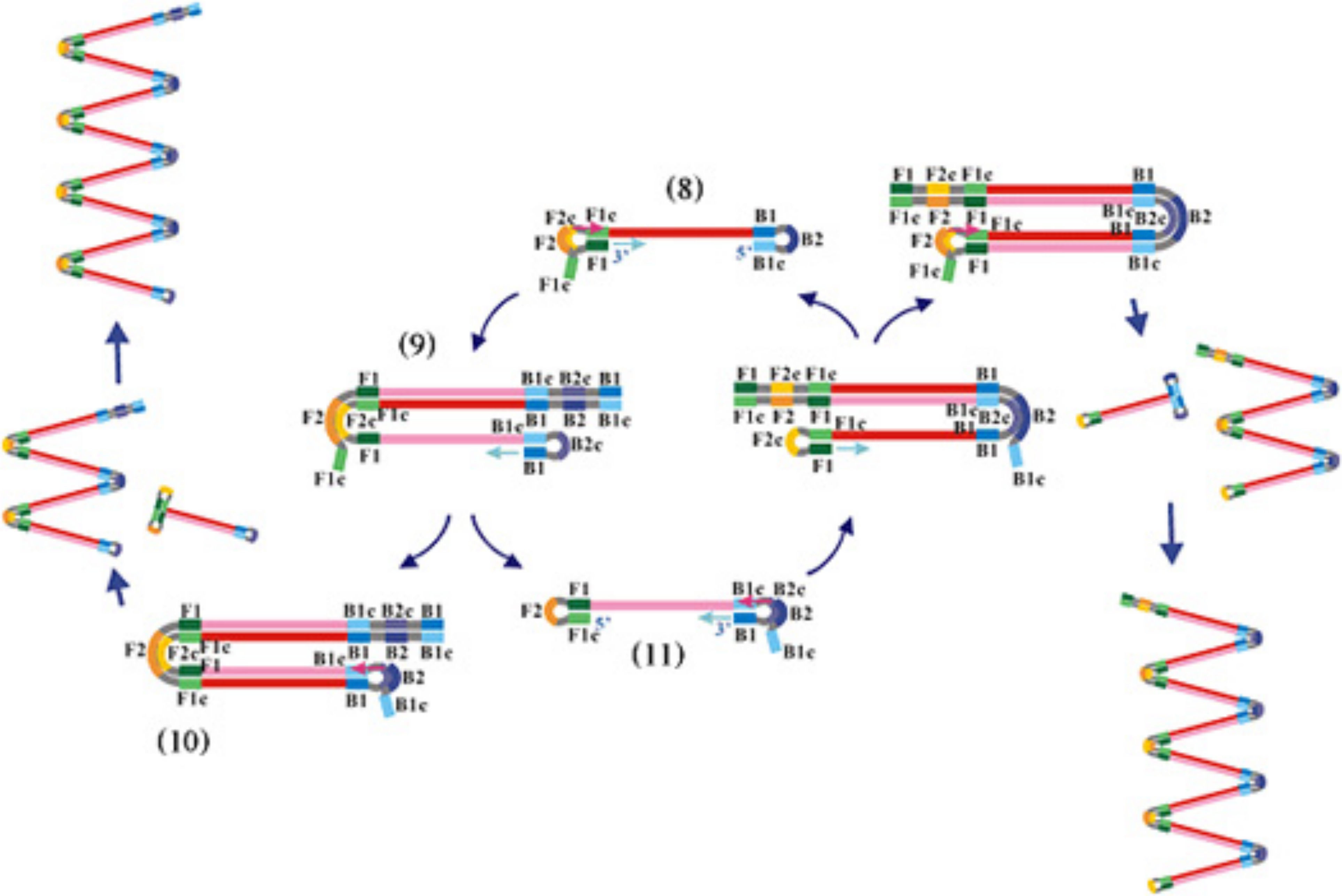


Dengue fever

- Transmitted by a bite of mosquito infected with dengue virus (genome size almost 11 kb)
- Febrile illness that affects infants, young children and adults with symptoms appearing 3-14 days after the infective bite.
- There are four serotypes (D1~ D4), whose genomes are about 70% identical one to each other.
- Second infection of the same serotype may cause severe symptoms; dengue hemorrhagic fever, abdominal pain, persistent vomiting, bleeding and breathing difficulty and is a potentially lethal.


DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
DEN-1 DEN-2 DEN-3 DEN-4
Jan. isoth.
10.C
Hawaii
Nepal
Bhutan
Hong Kong SAR
Macao SAR
Sudan
Galapagos islands
Madagascar - La Reunion
Easter island
Jul. isoth.
10.C
● Geographic extension of dengue, 2000-2008
🟧 Risk of dengue transmission
0
1,625
3,250
6,500 Kilometers

Template preparation using LAMP method



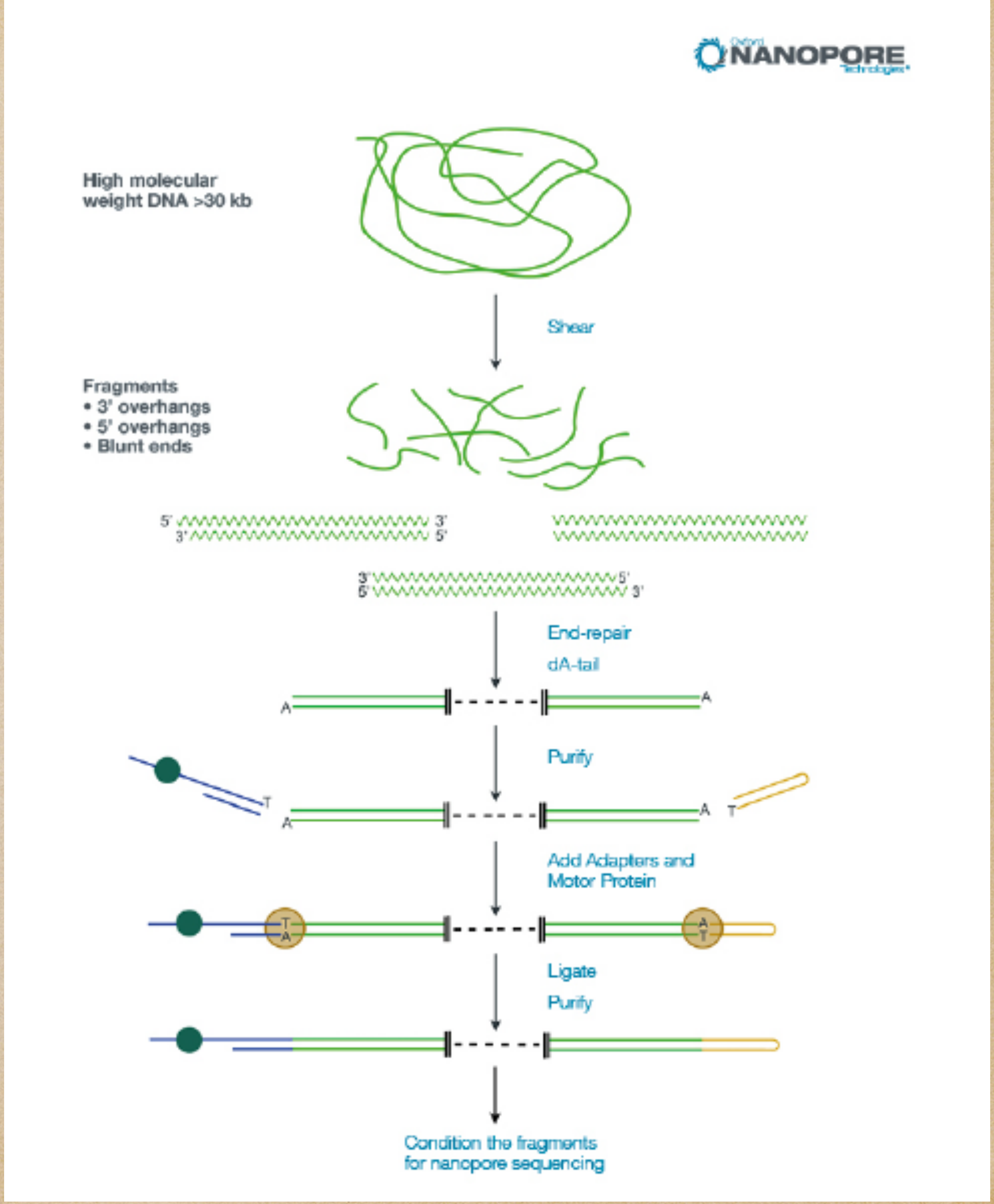
Sample preparation

LAMP Amplification

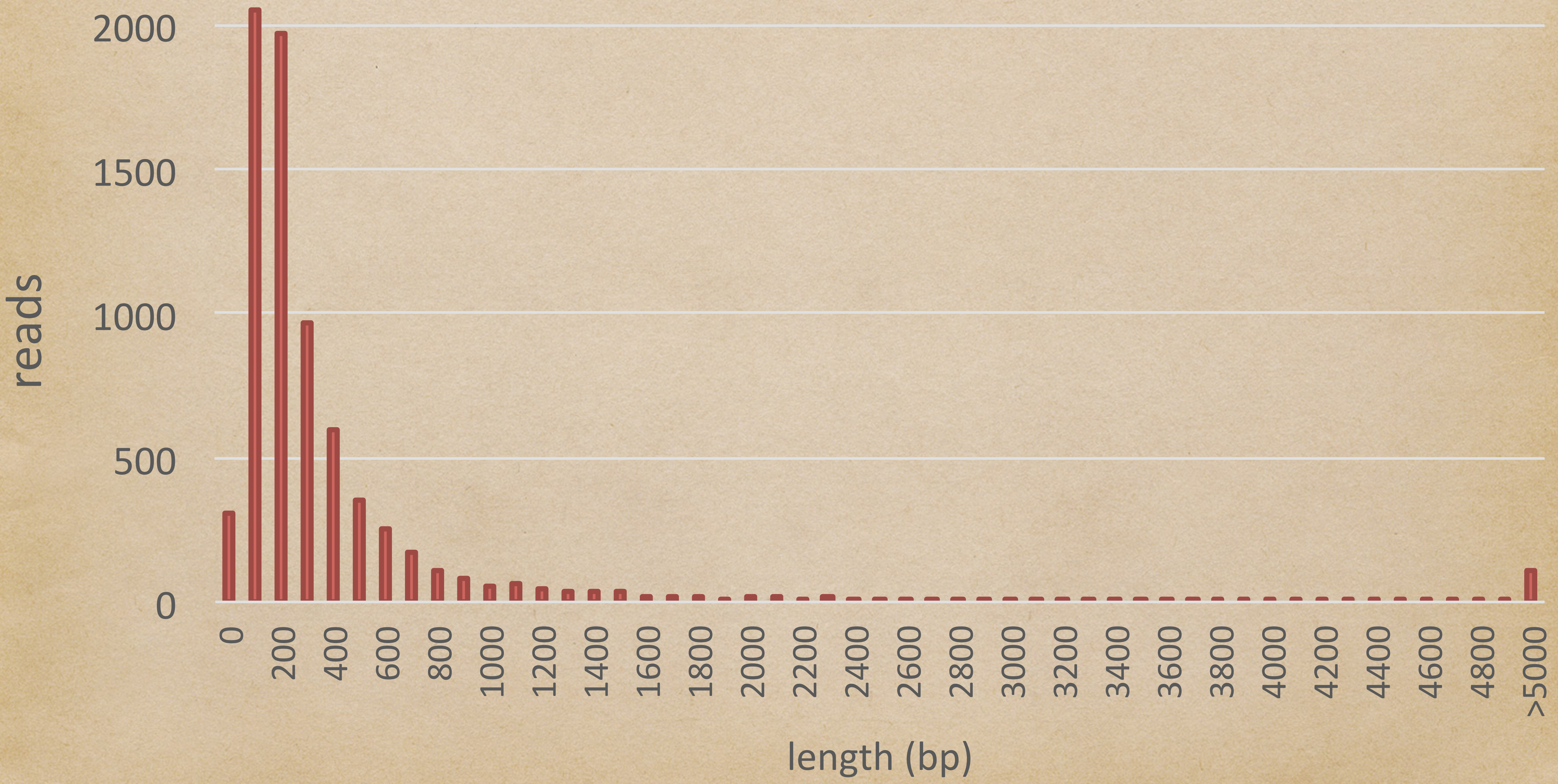
Serum (1 - 5 µL) -> Mix with Dry Lamp reagent kit -> 65 °C for 60 min -> Purification (AMPure) ->

Nonopore Sample Prep

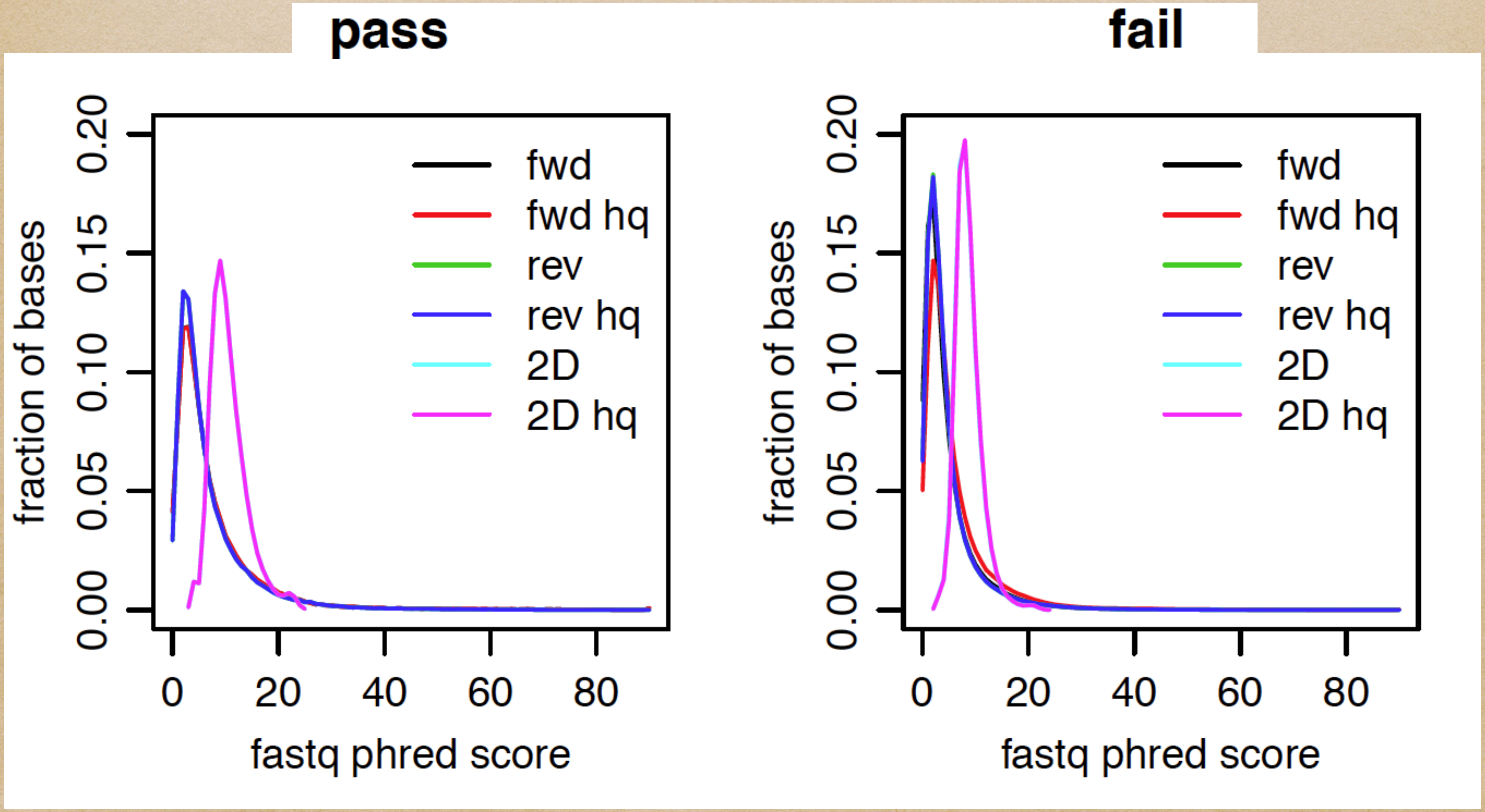
Easy, standard template preparation



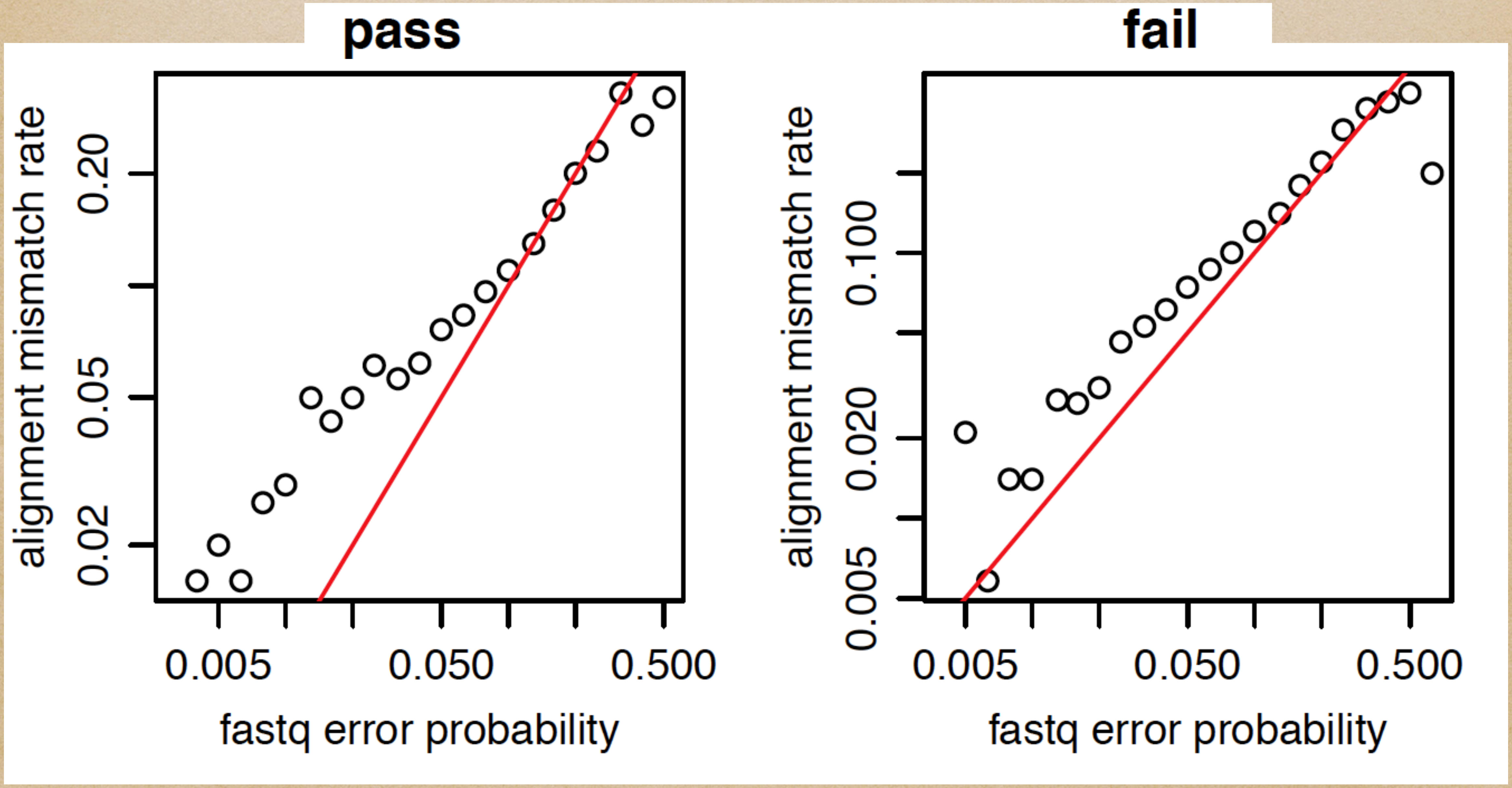
Distribution of read-length



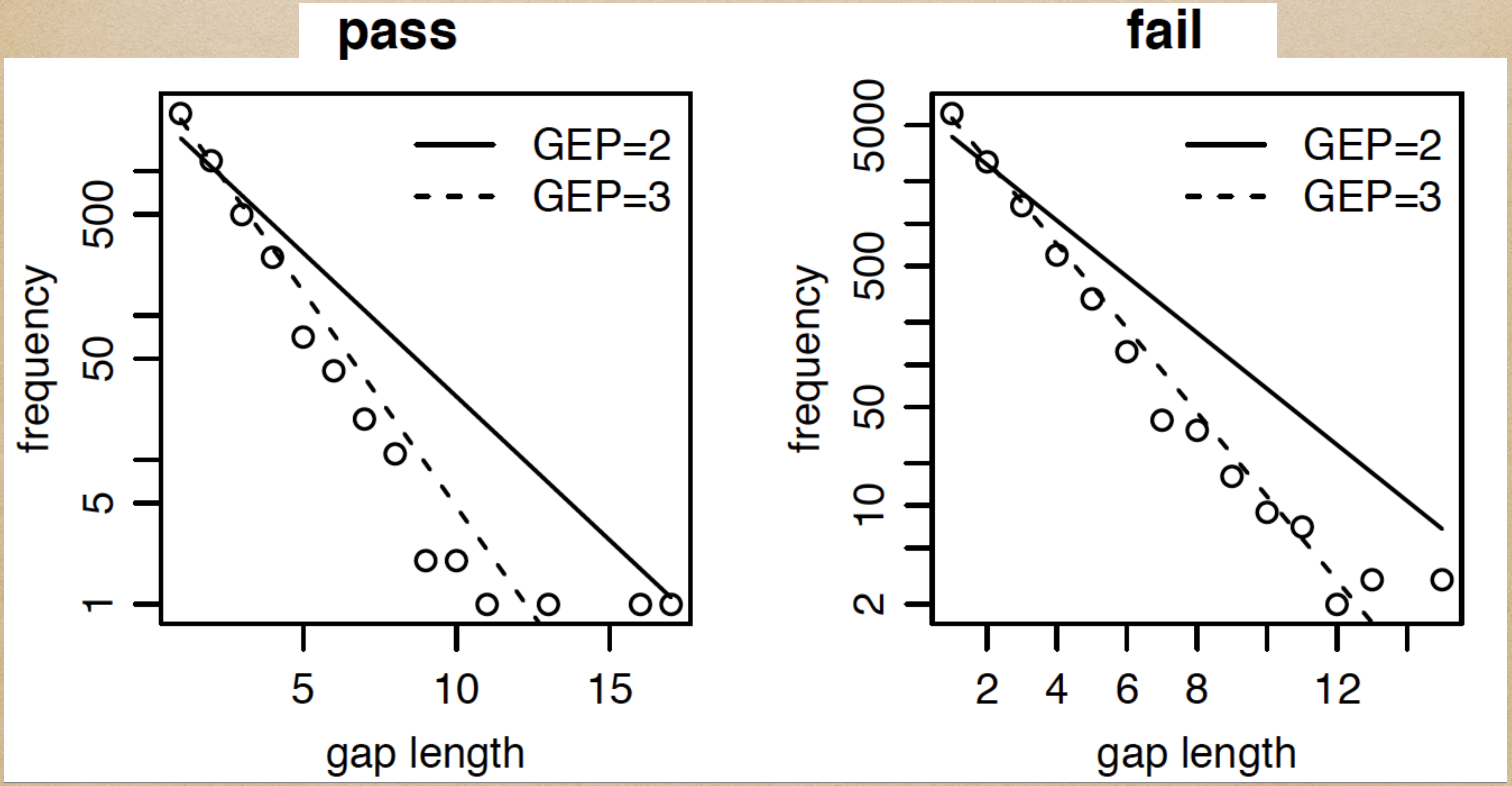
The distribution of fastq phred quality scores



The base error probabilities claimed by fastq agree quite well with the mismatch rates



The distribution of gap lengths



Best hit reads against Dengue genome with ssearch (D1 control)

Cut-off (E-val)	#Seq (D1)	#Seq (D2)	#Seq (D3)	#Seq (D4)	%Seq (D1)
1.0E-10	107	0	4	0	96%
1.0E-05	568	5	95	15	83%
All	1740	373	609	332	57%

Best hit reads against Dengue genome with ssearch (D2 control)

Cut-off (E-val)	#Seq (D1)	#Seq (D2)	#Seq (D3)	#Seq (D4)	%Seq (D1)
1.0E-10	1	630	7	6	98%
1.0E-05	168	4124	297	95	88%
All	4038	13989	4193	2728	56%

Best hit reads against Dengue genome with ssearch (unknown clinical sample)

Cut-off (E-val)	#Seq (D1)	#Seq (D2)	#Seq (D3)	#Seq (D4)	%Seq (D1)
1.0E-10	41	0	0	0	100%
1.0E-05	471	4	5	8	97%
All	3,301	1,398	1,372	1,372	44%

LAST alignments against Dengue genome (D2 control)

Sequence type	Total reads number	Number of alignments with error probability			
		<=0.1	<=0.01	<=0.001	<=0.0001
pass 2D	350	346	345	341	338
fail 2D	1492	889	807	744	681

With error probability <=00001 all reads aligned to D2

With error probability <=0001 just one read predicted to be chimeric

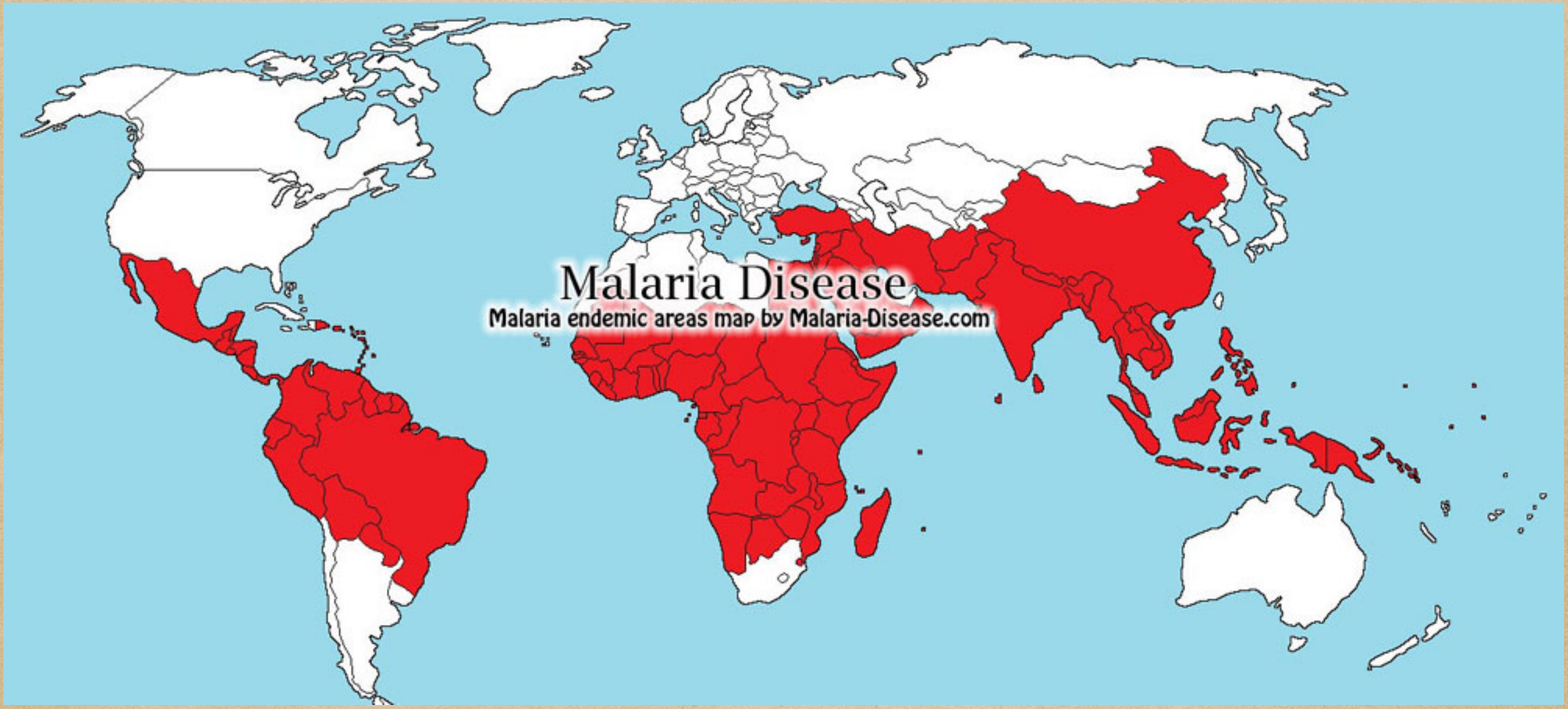
Conclusion

MinION, despite low sequencing accuracy, can be used for dengue virus serotyping

Still to come

Typing more clinical samples
Implementation of serotype prediction using LAST
Validation using Sanger/Illumina sequencing

Malaria parasite



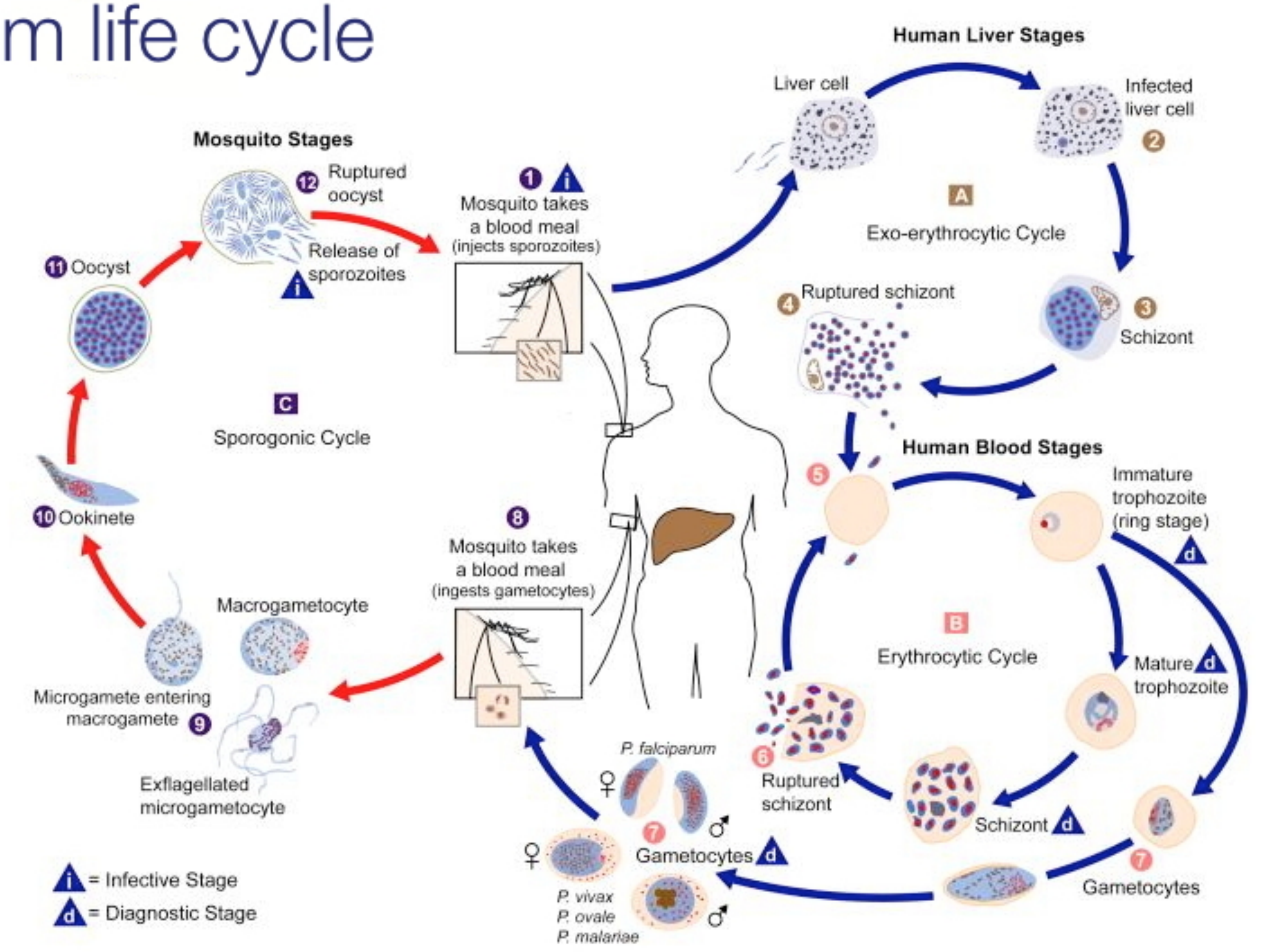
Malaria parasite

The Plasmodium life cycle

The sexual stage of the cycle is completed in the mosquito, the asexual one in human red blood cells.

Image from the Laboratory Identification of Parasites of Public Health Concern (DPDx) of the USA government of Centers for Disease Control and Prevention (CDC)

(http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/Malaria_il.htm)



Emergence of Drug Resistant Mutations

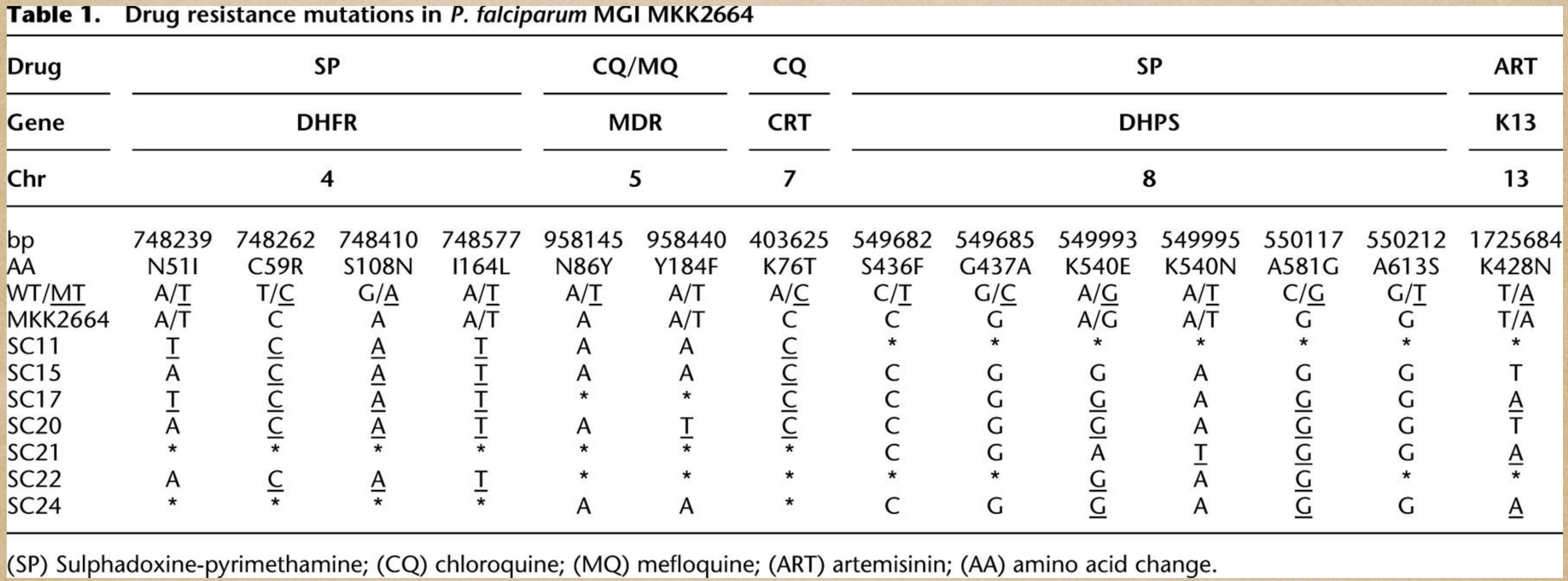


Table 1. Drug resistance mutations in *P. falciparum* MGI MKK2664

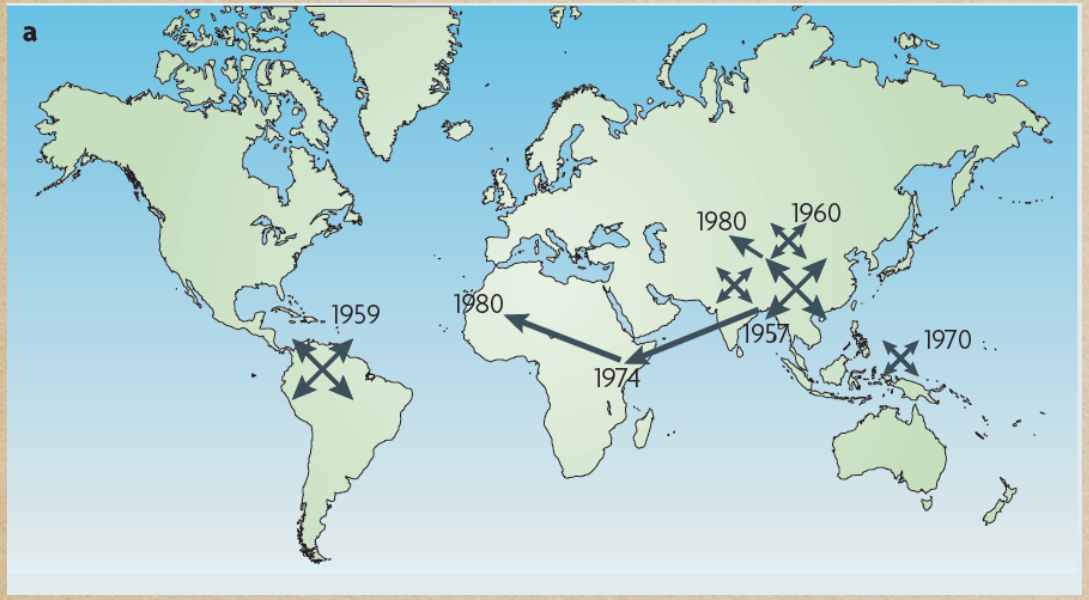
Drug	**SP**				**CQ/MQ**		**CQ**	**SP**							**ART**
Gene	**DHFR**				**MDR**		**CRT**	**DHPS**							**K13**
Chr	**4**				**5**		**7**	**8**							**13**
bp	748239	748262	748410	748577	958145	958440	403625	549682	549685	549993	549995	550117	550212	1725684
AA	N51I	C59R	S108N	I164L	N86Y	Y184F	K76T	S436F	G437A	K540E	K540N	A581G	A613S	K428N
WT/MT	A/T	T/C	G/A	A/T	A/T	A/T	A/C	C/T	G/C	A/G	A/T	C/G	G/T	T/A
MKK2664	A/T	C	A	A/T	A	A/T	C	C	G	A/G	A/T	G	G	T/A
SC11	T	C	A	T	A	A	C	*	*	*	*	*	*	*
SC15	A	C	A	T	A	A	C	C	G	G	A	G	G	T
SC17	T	C	A	T	*	*	C	C	G	G	A	G	G	A
SC20	A	C	A	T	A	T	C	C	G	G	A	G	G	T
SC21	*	*	*	*	*	*	*	C	G	A	T	G	G	A
SC22	A	C	A	T	*	*	*	*	*	G	A	G	*	*
SC24	*	*	*	*	A	A	*	C	G	G	A	G	G	A

(SP) Sulphadoxine-pyrimethamine; (CQ) chloroquine; (MQ) mefloquine; (ART) artemisinin; (AA) amino acid change.

Nair at al. (2014) Single-cell genomics for dissection of complex malaria infections. Genome Res. 24(6):1028-38.

Spreading drug resistance

Chloroquine resistance

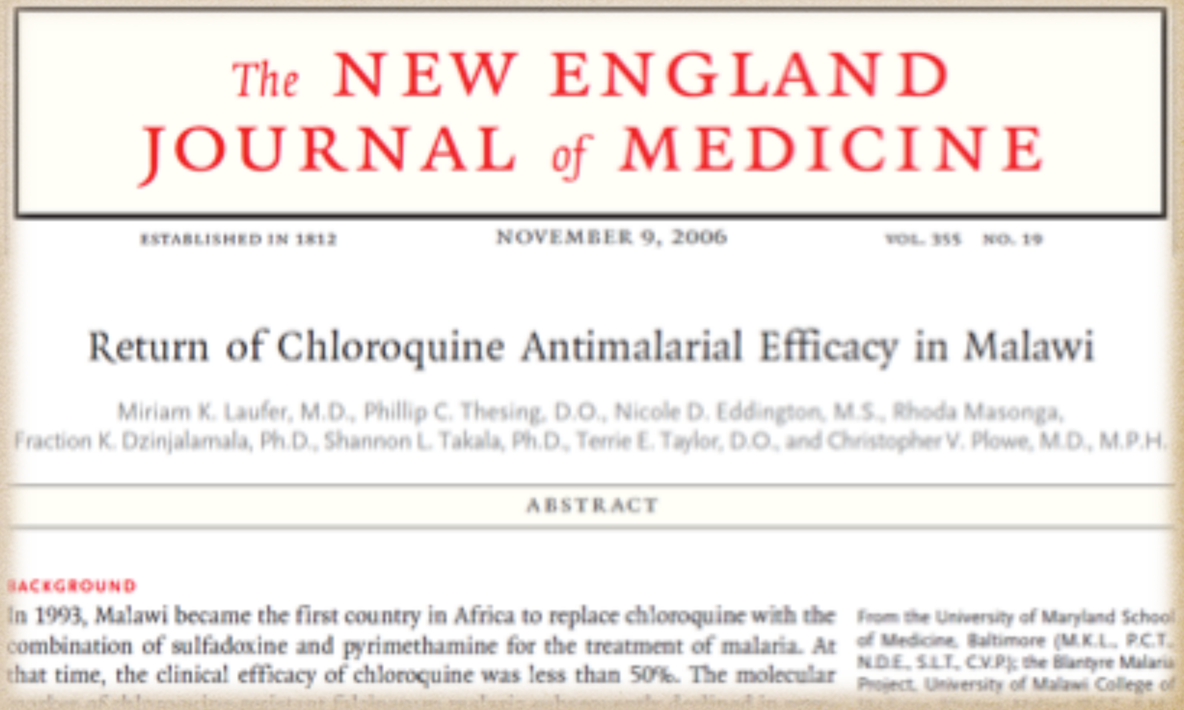


Sulfadoxine resistance

Dondorp et al. Nat. Microbiol (2010)

Declining of resistance parasite population

A study in Malawi, reported that population of CQ resistant P. falciparum (CQR) has decreased.



The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 NOVEMBER 9, 2006 VOL. 355 NO. 19

Return of Chloroquine Antimalarial Efficacy in Malawi

Miriam K. Laufer, M.D., Phillip C. Thesing, D.O., Nicole D. Eddington, M.S., Rhoda Masonga, Fraction K. Dzinjalamala, Ph.D., Shannon L. Takala, Ph.D., Terrie E. Taylor, D.O., and Christopher V. Plowe, M.D., M.P.H.

ABSTRACT

ACKGROUND

In 1993, Malawi became the first country in Africa to replace chloroquine with the combination of sulfadoxine and pyrimethamine for the treatment of malaria. At that time, the clinical efficacy of chloroquine was less than 50%. The molecular

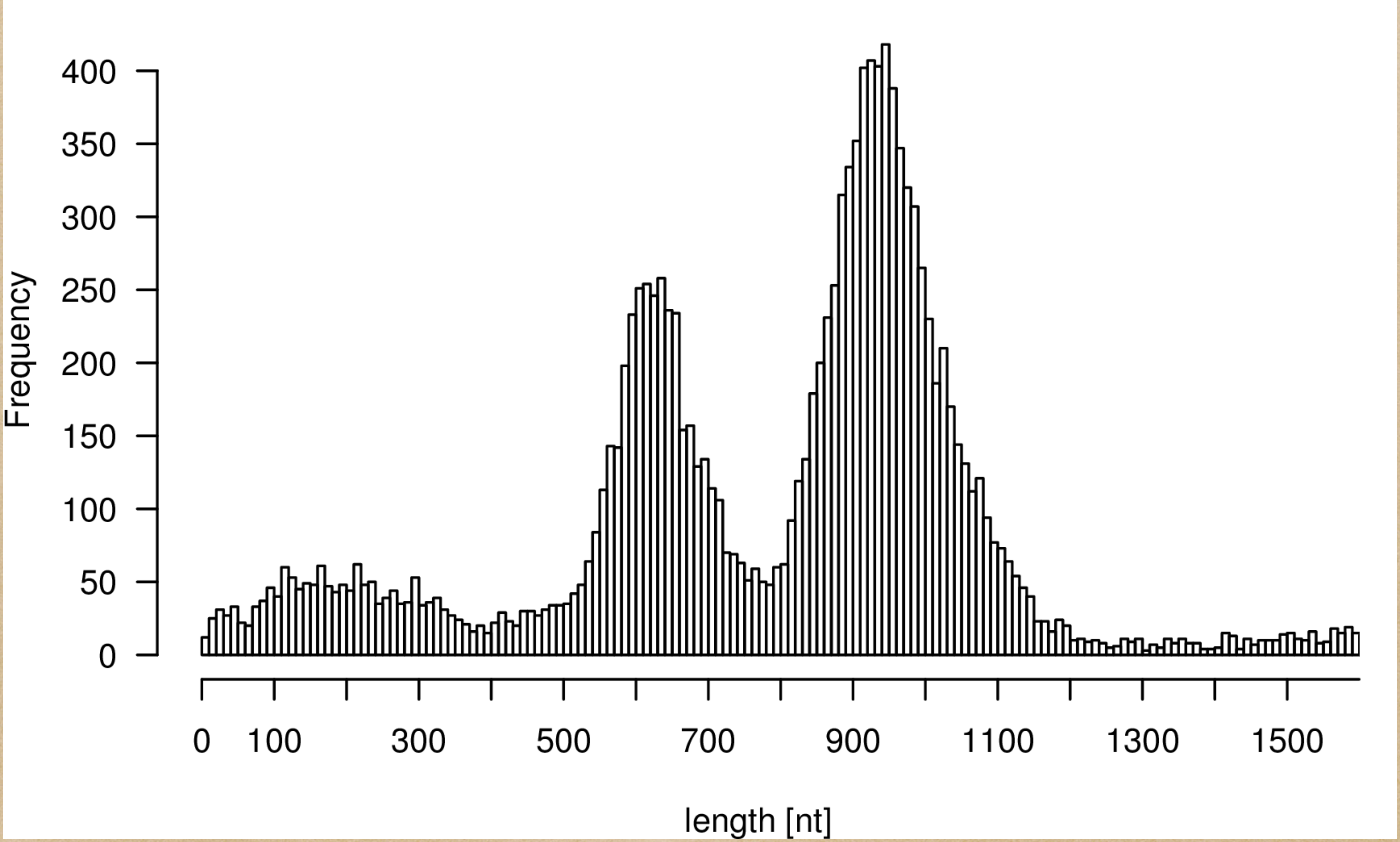
From the University of Maryland Schoo of Medicine, Baltimore (M.K.L., P.C.T., N.D.E., S.L.T., C.V.P.); the Blantyre Malaria Project, University of Malawi College of

Declining of 76T mutation also reported in some other countries (i.e. Benin, Kenya, and Senegal) though at a slower rate.

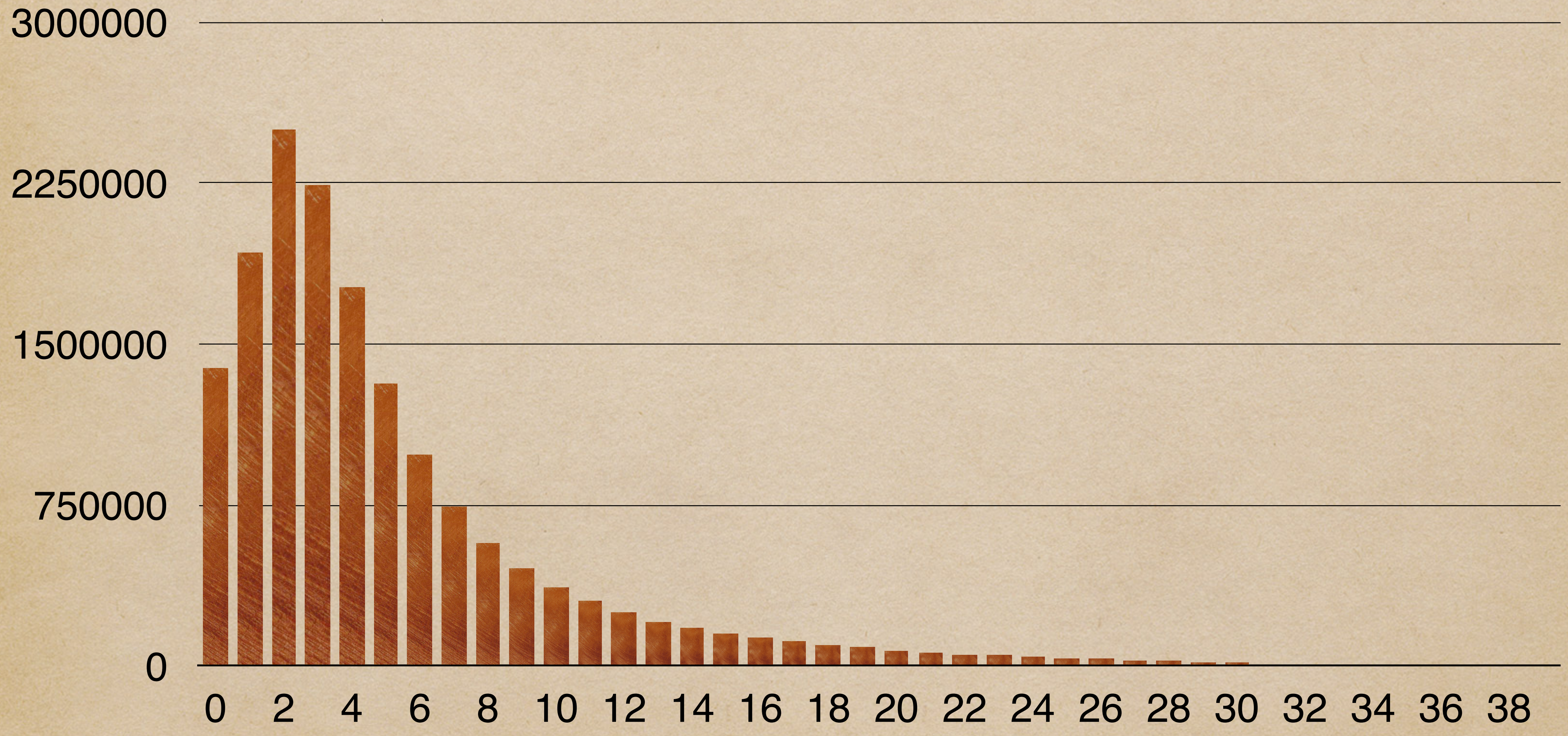
Targets for genotyping

Gene definition	Chromosome	Start Position	End position	strand	length
apocytochrome_b	>M76611	3438	4680	+	1243
PfATPase6_1	>Pf3D7_01_v3	267134	269239	-	2106
PfATPase6_2	>Pf3D7_01_v3	264641	267470	+	2830
pfmrp1_1	>Pf3D7_01_v3	464622	467289	+	2668
pfmrp1_2	>Pf3D7_01_v3	466960	470216	-	3257
DHFR-TS	>Pf3D7_04_v3	747923	749956	+	2034
PfTCTP	>Pf3D7_05_v3	467406	468316	-	911
pfmdr1	>Pf3D7_05_v3	957756	962218	+	4463
PfCRT_1	>Pf3D7_07_v3	403089	404828	+	1740
PfCRT_2	>Pf3D7_07_v3	404757	406466	-	1710
DHPS	>Pf3D7_08_v3	548039	550780	+	2742
ABC_transporter	>Pf3D7_08_v3	670708	675573	-	4866
K13-propeller	>Pf3D7_13_v3	1724572	1727035	-	2464

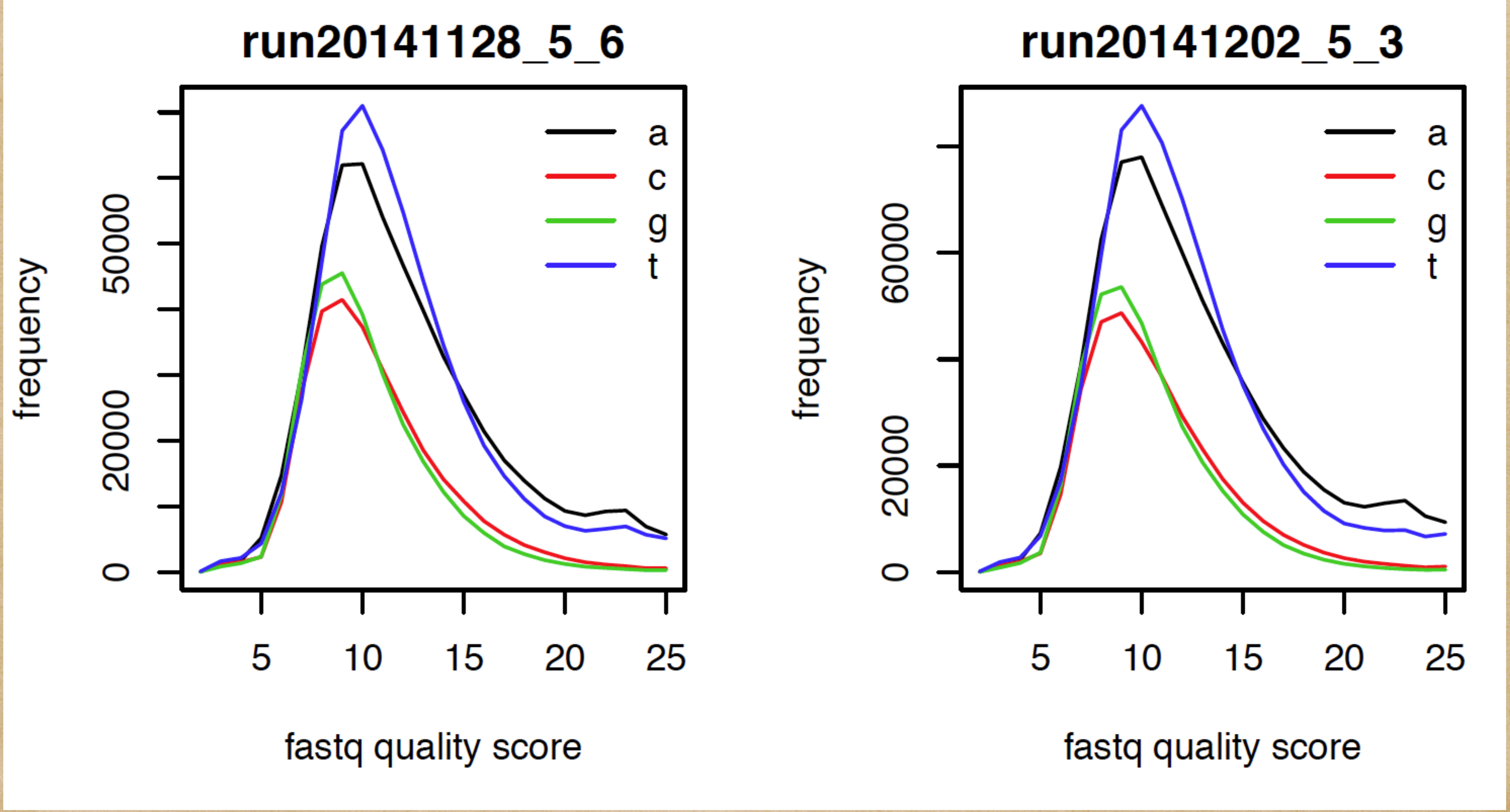
Distribution of reads lengths first sequencing (summer 2014)



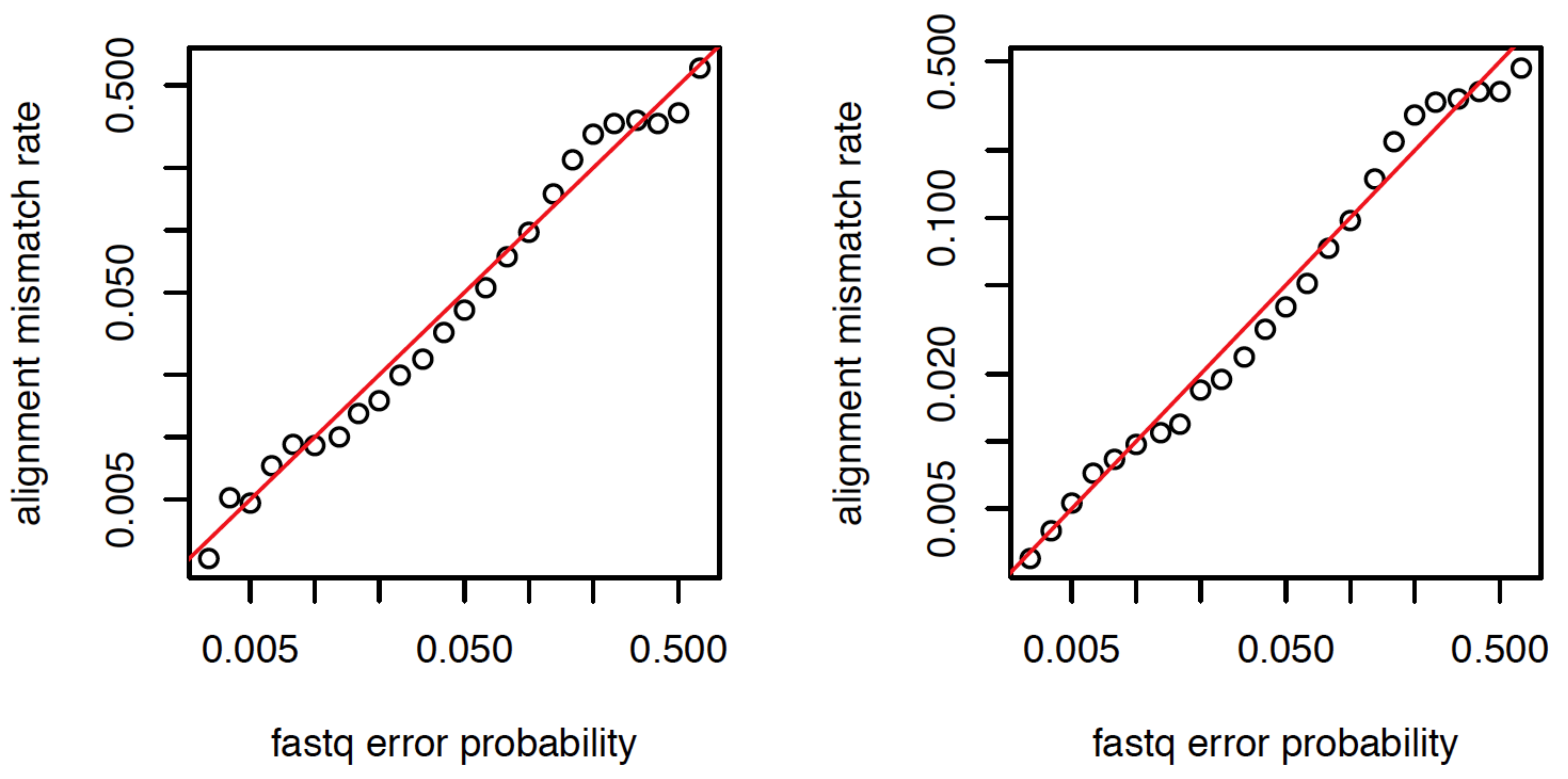
Quality of nanopore reads



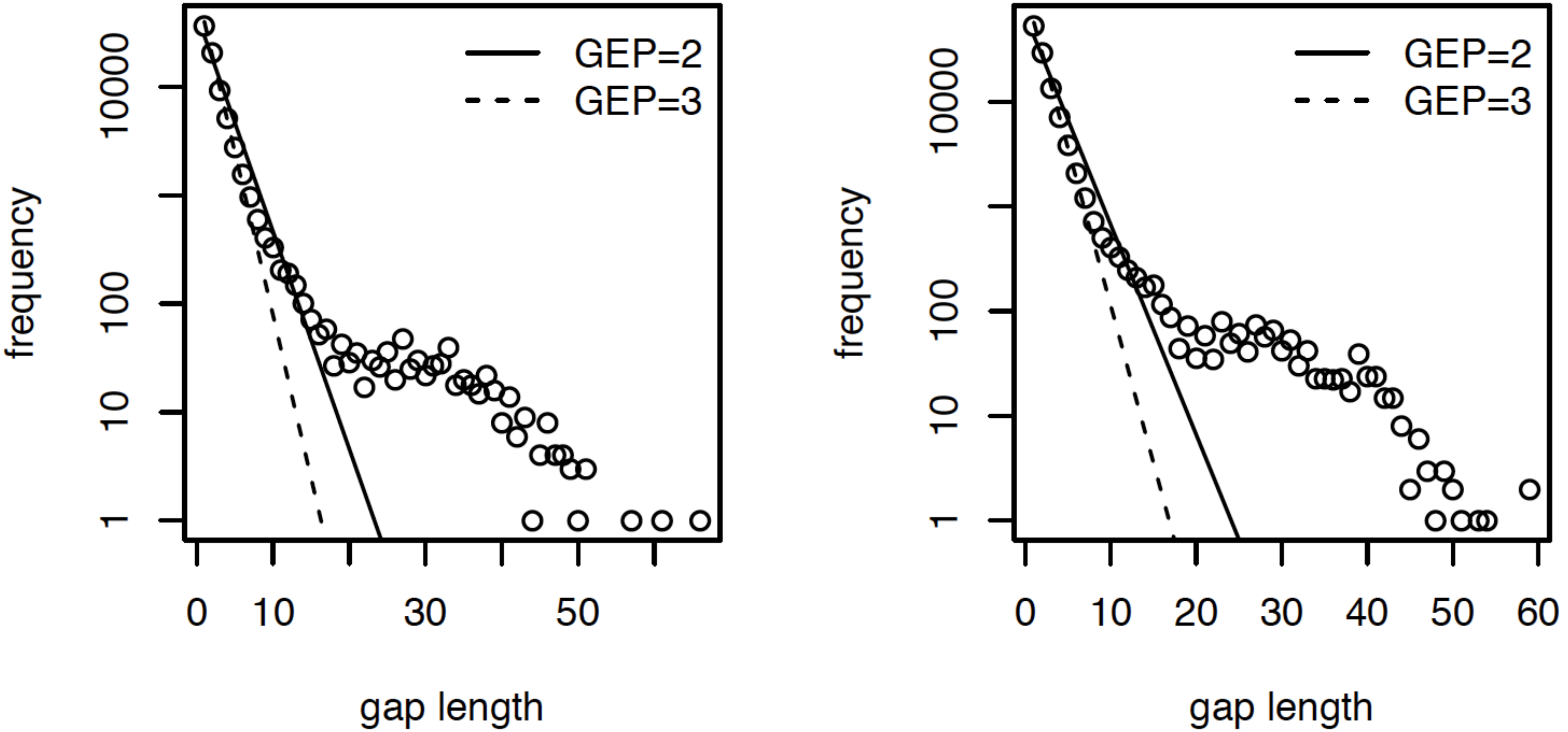
Bases annotated with high quality scores are AT-rich



The fastq error probabilities agree well with the mismatch rates

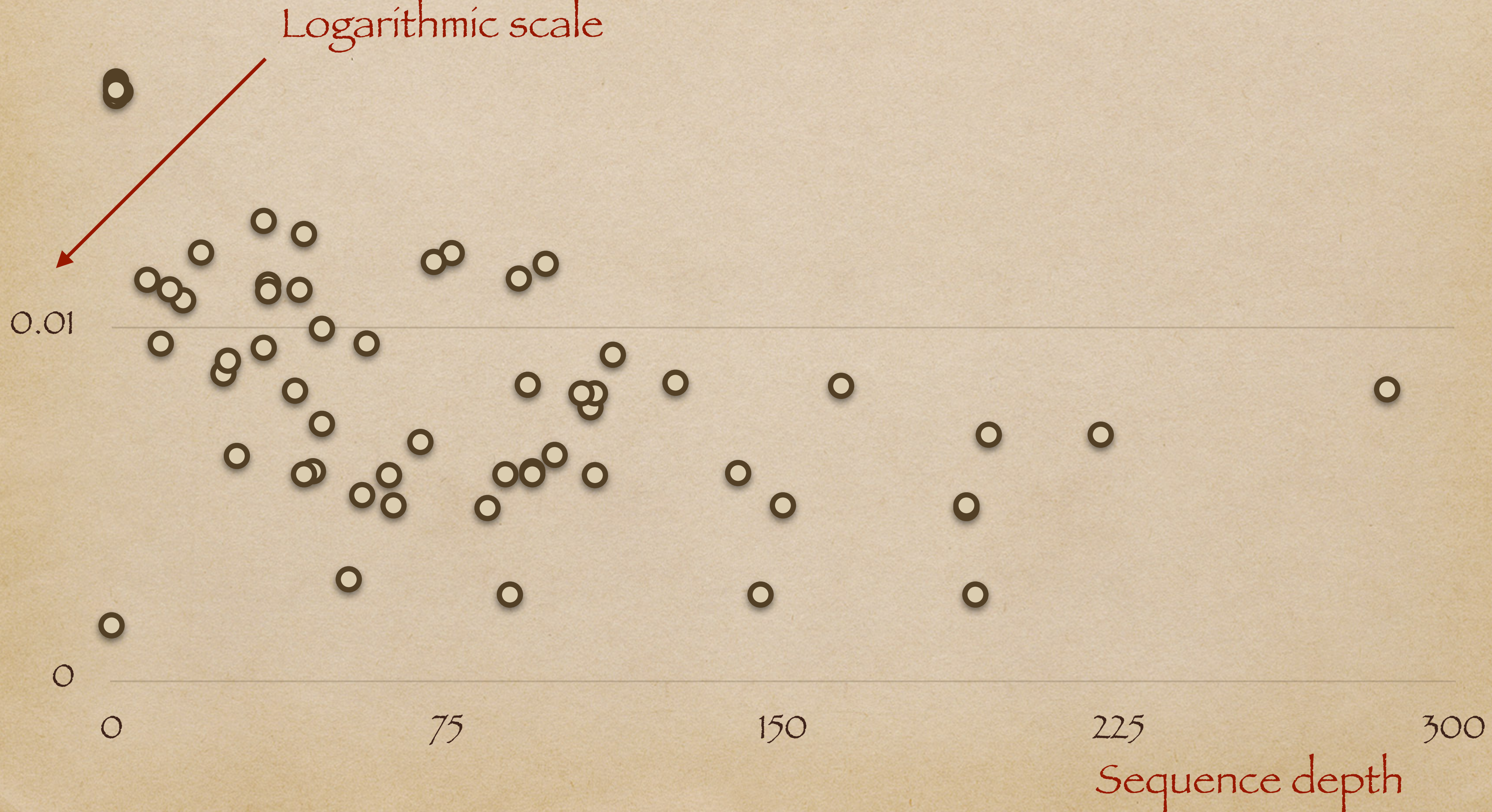


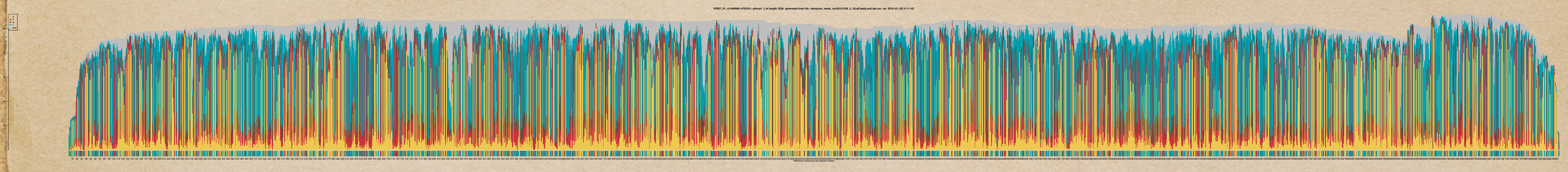
The distribution of gap lengths in alignments to the target sequences

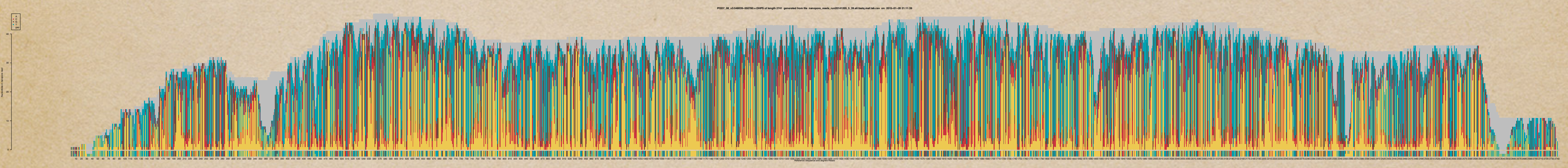


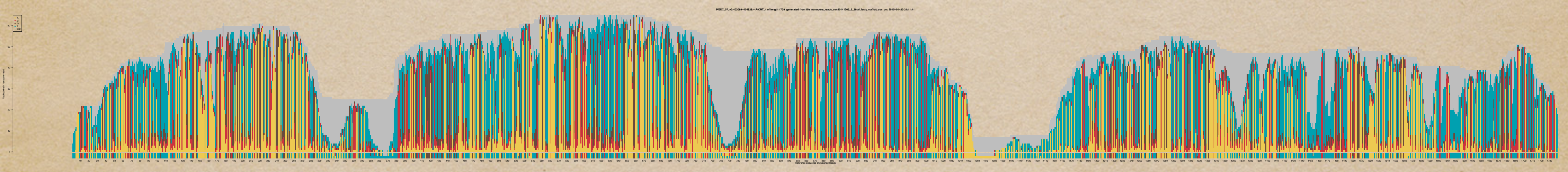
Miscall rate against amplicon length

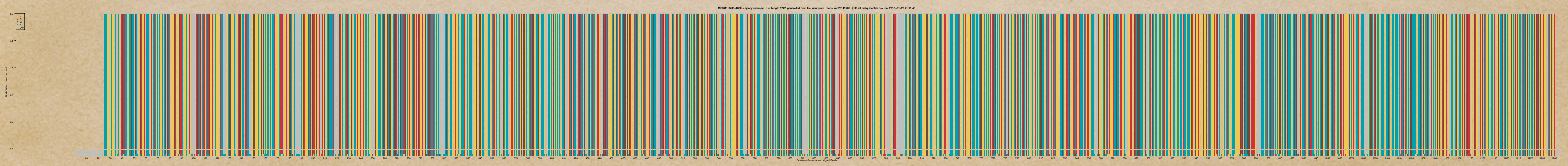
Miscall rate against sequencing depth



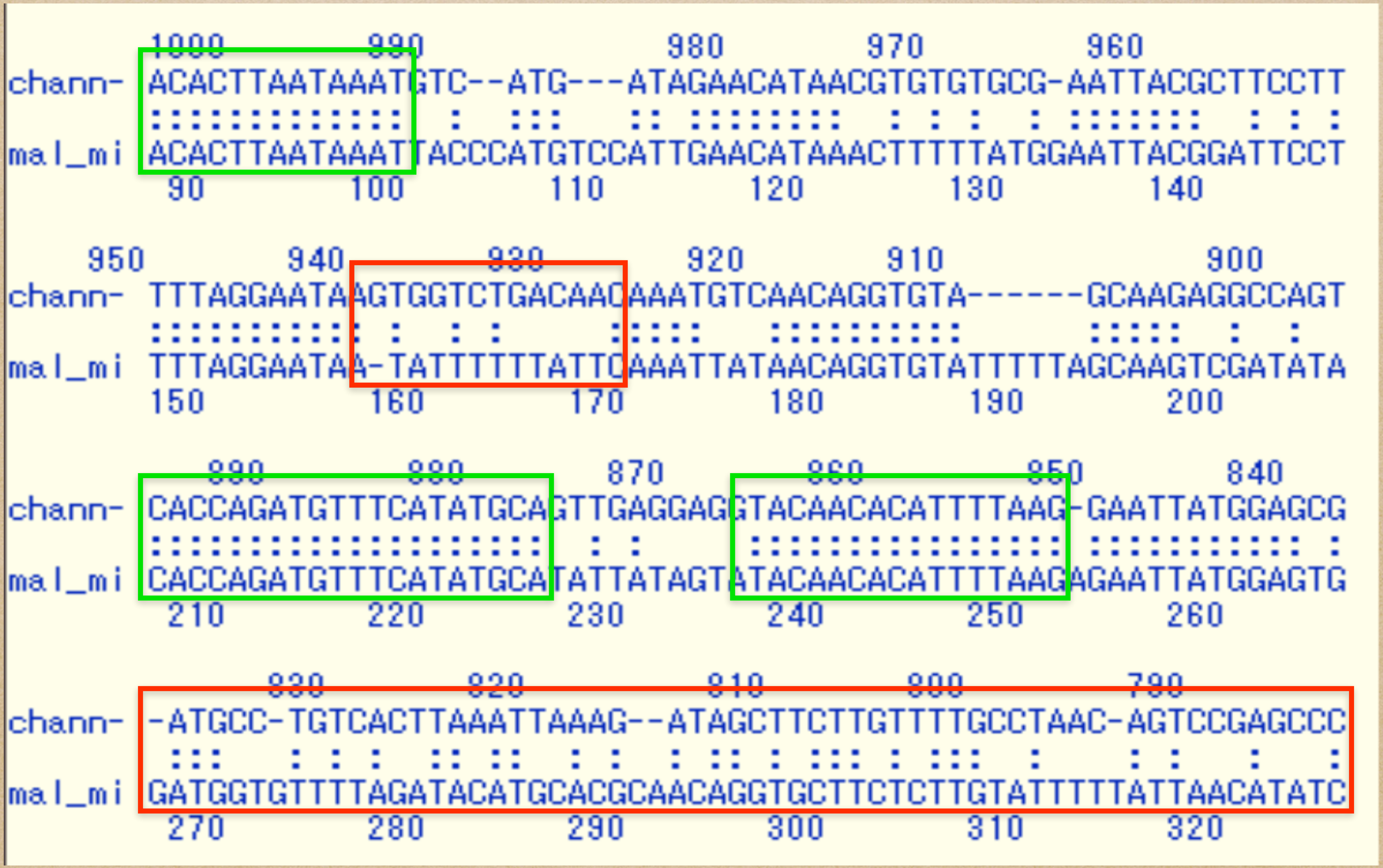

Pf3D7_01_v3:466960–470216:–:pfmrp1_2 of length 3256 generated from file nanopore_reads_run20141205_5_39.all.fastq.maf.tab.csv on: 2015–01–20 21:11:42


Pf3D7_08_v3:548039–550780:+:DHPS of length 2741 generated from file nanopore_reads_run20141205_5_39.all.fastq.maf.tab.csv on: 2015–01–20 21:11:39


Pf3D7_07_v3:403089–404828:+:PfCRT_1 of length 1739 generated from file nanopore_reads_run20141205_5_39.all.fastq.maf.tab.csv on: 2015–01–20 21:11:41


M76611:3438–4680:+:apocytochrome_b of length 1242 generated from file nanopore_reads_run20141205_5_39.all.fastq.maf.tab.csv on: 2015–01–20 21:11:43

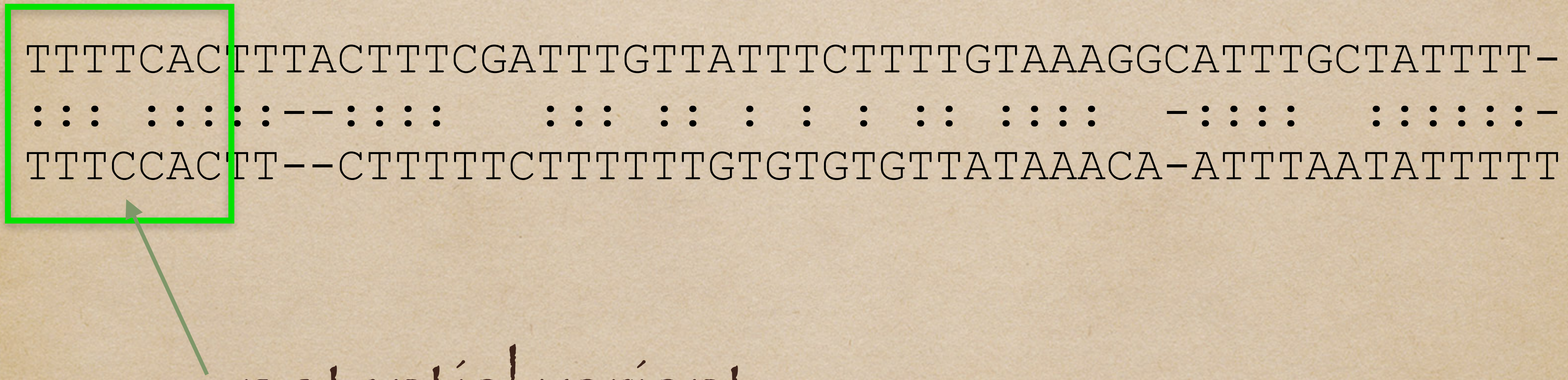
ssearch alignments



```
       1000       990            980       970        960
chann- ACACTTAATAAATGTC--ATG---ATAGAACATAACGTGTGTGCG-AATTACGCTTCCTT
       :::::::::::::  :  :::   :: ::::::::  : : :  : :::::::  : : :
mal_mi ACACTTAATAAATTACCCATGTCCATTGAACATAAACTTTTTATGGAATTACGGATTCCT
        90       100       110       120       130       140

    950       940       930       920       910             900
chann- TTTAGGAATAAGTGGTCTGACAACAAATGTCAACAGGTGTA------GCAAGAGGCCAGT
       ::::::::::: :  : :     :::::   ::::::::::      :::::  :  :
mal_mi TTTAGGAATAA-TATTTTTTATTCAAATTATAACAGGTGTATTTTTAGCAAGTCGATATA
       150        160       170       180       190       200

          890       880       870       860        850       840
chann- CACCAGATGTTTCATATGCAGTTGAGGAGGTACAACACATTTTAAG-GAATTATGGAGCG
       ::::::::::::::::::::  : :     :::::::::::::::: ::::::::::: :
mal_mi CACCAGATGTTTCATATGCATATTATAGTATACAACACATTTTAAGAGAATTATGGAGTG
        210       220       230       240       250       260

             830       820         810       800        790
chann- -ATGCC-TGTCACTTAAATTAAAG--ATAGCTTCTTGTTTTGCCTAAC-AGTCCGAGCCC
        :::   : : :  :: ::  : :  : :: : ::: : :::  :    : :   :   :
mal_mi GATGGTGTTTTAGATACATGCACGCAACAGGTGCTTCTCTTGTATTTTTATTAACATATC
        270       280       290       300       310       320
```

3+1+3 rule

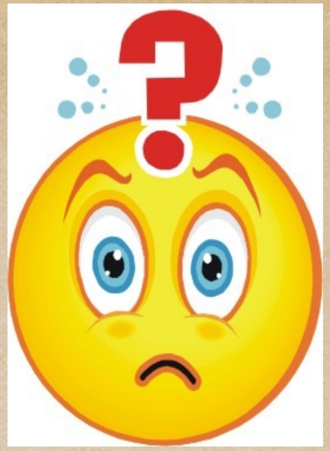
1. Smith-Waterman alignment
2. Sliding window of seven nt for base call
3. 3+1+3 rule applied



```
TTTTCACTTTACTTTCGATTTGTTATTTCTTTTGTAAAGGCATTTGCTATTTT-
::: :::::--::::   ::: :: : : : :: ::::  -::::  ::::::-
TTTCCACTT--CTTTTTCTTTTTTGTGTGTGTTATAAACA-ATTTAATATTTTT
```

potential variant

Conclusions

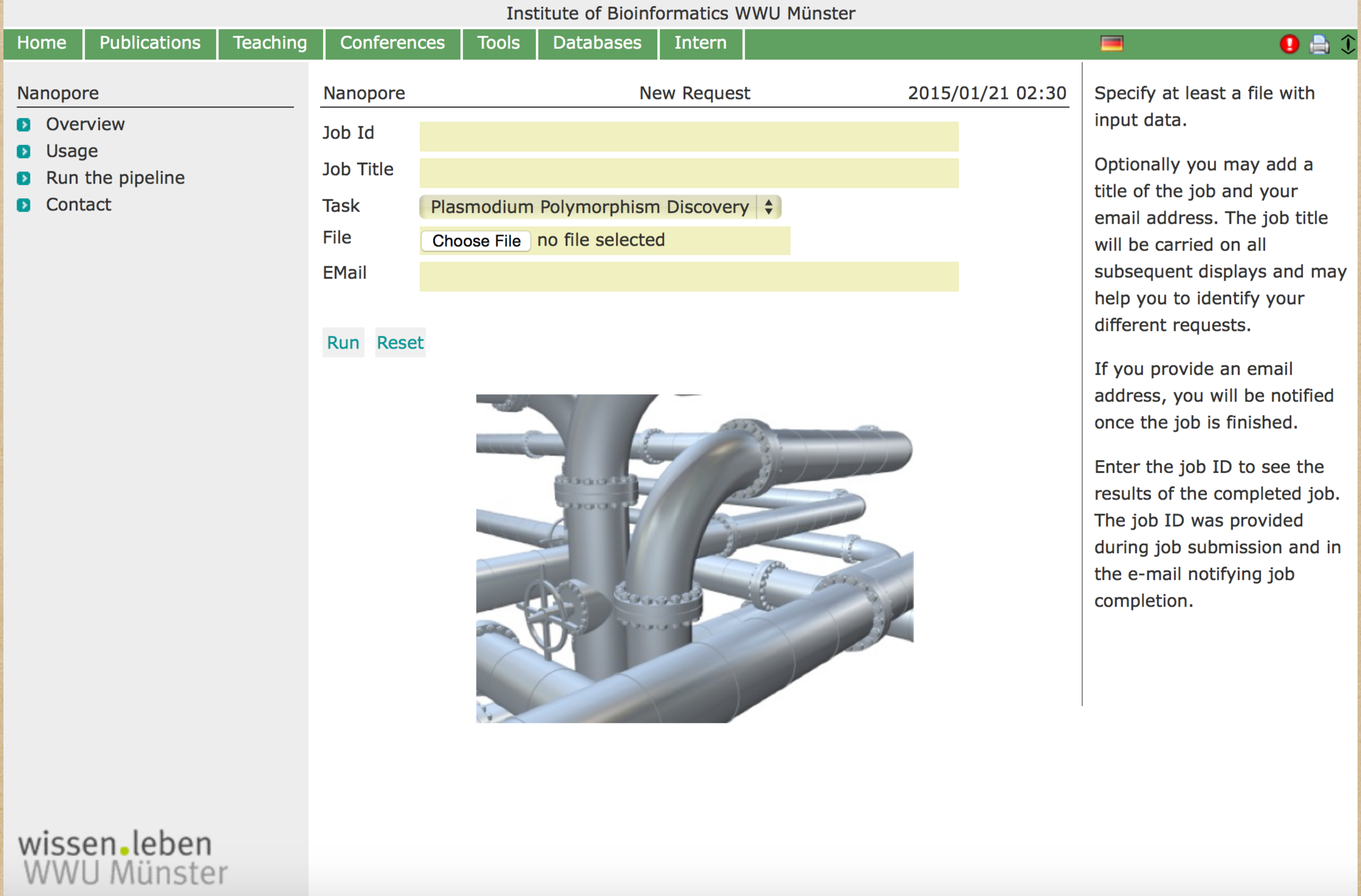


1. Mutation calls for “known” SNVs in plasmodium with the accuracy of up to 85%
2. At the moment not good for indels call
3. Needs for indexing (bar-coding) method for cost-effective field applications

However…

It opens interesting options in the field studies

Nanopore pipeline



Nanopore New Request 2015/01/21 02:30

Job Id

Job Title

Task: Plasmodium Polymorphism Discovery

File: Choose File no file selected

EMail

Specify at least a file with input data.

Optionally you may add a title of the job and your email address. The job title will be carried on all subsequent displays and may help you to identify your different requests.

If you provide an email address, you will be notified once the job is finished.

Enter the job ID to see the results of the completed job. The job ID was provided during job submission and in the e-mail notifying job completion.

Data upload to the server via web interface

Sequence conversion from fast5 to fast

Pairwise sequence alignment to the reference sequence using LAST or ssearch

Nucleotide count at each position of the reference sequence and consensus call based on majority rule

SNP call (plasmodium) or serotype classification (dengue)

WWW report

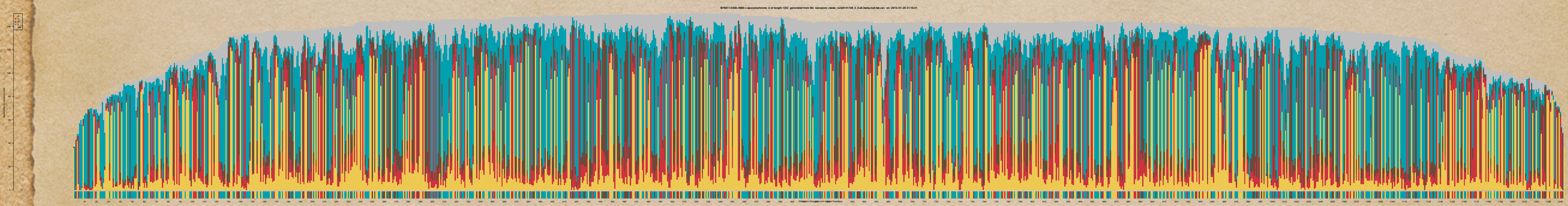
WWW report

Short textual summary

For plasmodium: a table listing potential polymorphisms

For dengue: a classification decision

Graphical presentation of the sequencing results



Indonesia
Sample collection

Josef Tuda
Arthur E. Mongan

Japan
Sequencing + bioinformatics

Yutaka Suzuki
Junya Yamagishi
Mihoko Imada
Keiko Toya
Hiroyuki Wakaguri
Kiyomi Imamura
Kazumi Abe

Martin Frith

Germany
Bioinformatics

Tabea Kischka
Norbert Grundman