

# Many-To-One-To-Many: Pooling and Demultiplexing

Dr. Armin Töpfer

2018-06-14

## SEQUEL SYSTEM PERFORMANCE



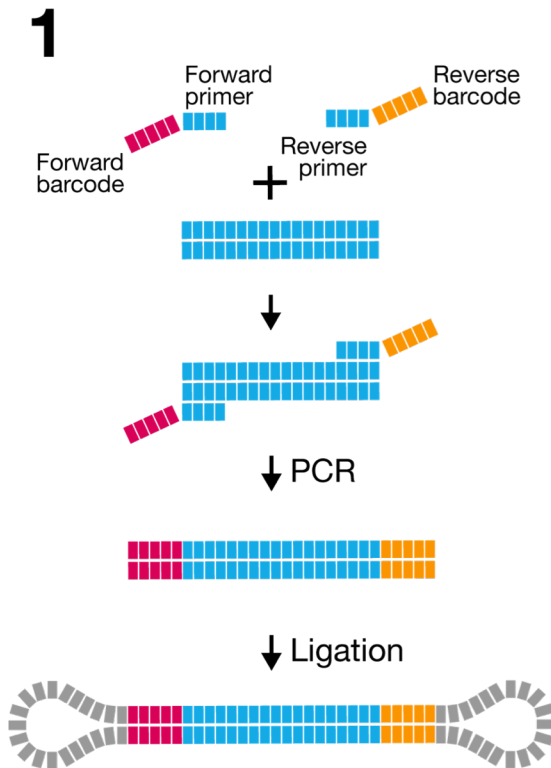
### 1M ZMWs enables

- ❖ 2.5-fold human
- ❖ > 8 microbial assemblies
- ❖ 8 minor variant samples for 1% accuracy
- ❖ >100 small amplicons

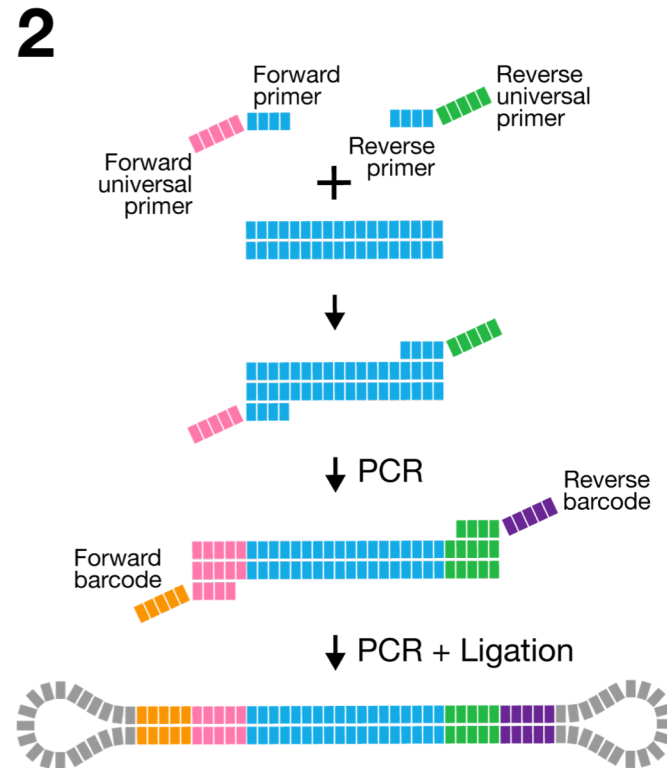
Barcode > Pool > Sequence > Demultiplex > Analyze



# WORKFLOW

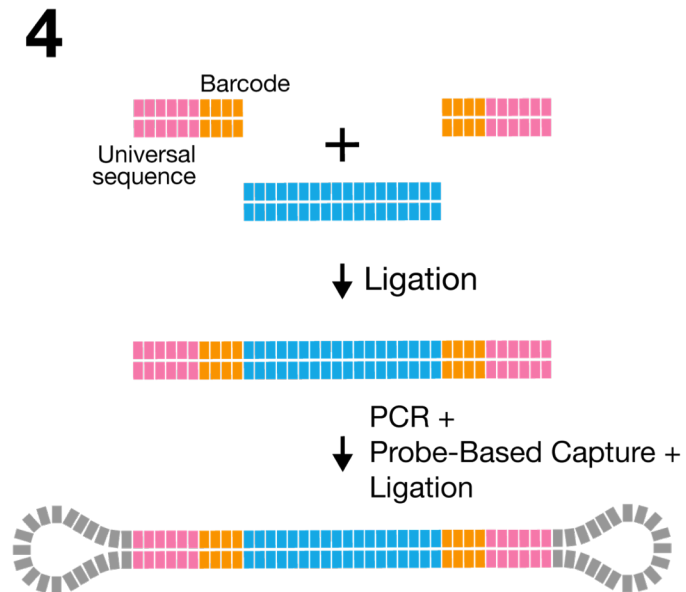
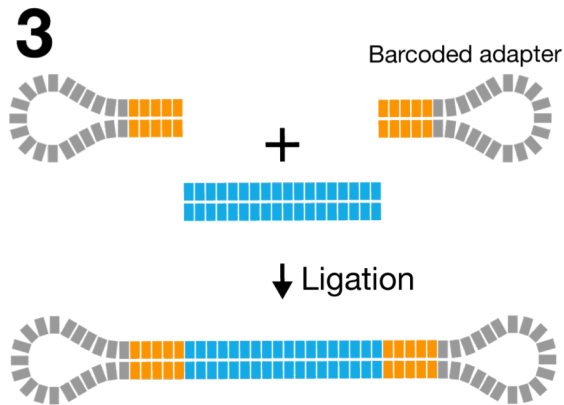


384 barcodes for  
sequence specific primers



96 barcoded universal primers

# WORKFLOW

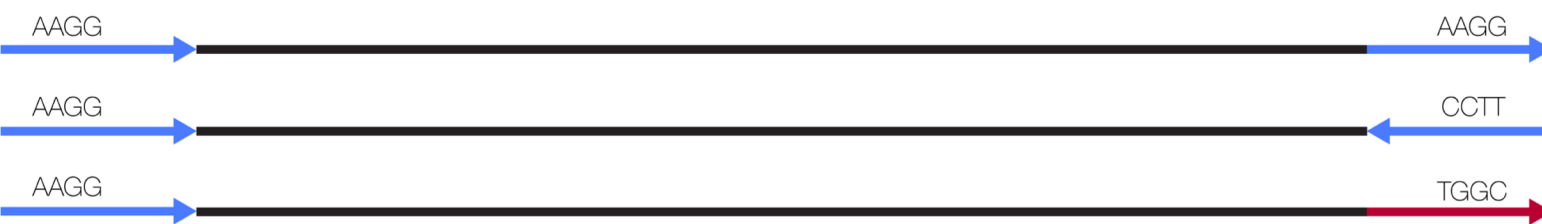
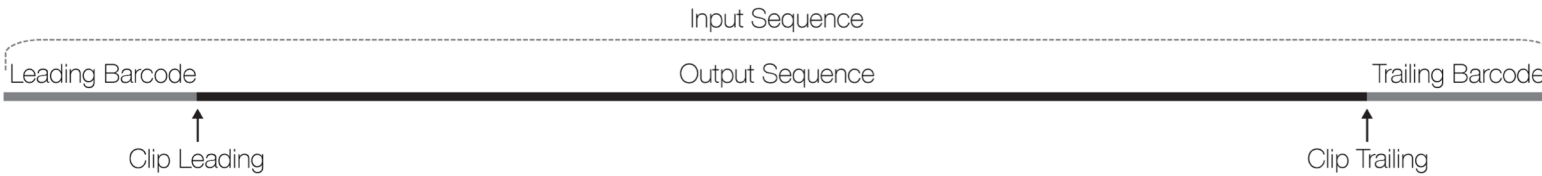


96 barcoded adapters

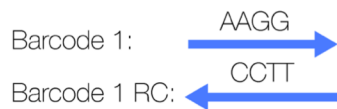
2 x 8 microbial barcoded adapters



Do not worry about tailed or symmetric.  
 Call it **SAME** or **DIFFERENT** barcodes in a pair.



Terminology	
library	lima
tailed	same
symmetric	same
asymmetric	different





- Same insert size → equimolar pooling
- Different insert size → weighted pooling



- Demultiplexing powered by *lima* <https://github.com/pacificbiosciences/barcoding>
- Easy configuration via SMRT Link

▼ Barcoded Sample Options

Sample is Barcoded  YES  NO

Barcode Set  ▼

Same Barcodes on Both Ends of Sequence  ON  OFF ⓘ

Barcode Data ⓘ

Barcoded Sample Name File ⓘ

- Label your bio samples

	A	B	C	D	E	F	G	H
1	Barcode Name	Bio Sample Name (allowed characters: alphanumeric; space; dot; underscore; hyphen)						
2	bc1001--bc1001	ecoli						
3	bc1002--bc1002	yeast						
4	bc1003--bc1003	hiv-env						
5	bc1004--bc1004	hiv-patient1						
6	bc1005--bc1005	hiv-patient2						
7	bc1006--bc1006	bc1006--bc1006						
8	bc1007--bc1007	bc1007--bc1007						



Barcode

Pool

Sequence

Demultiplex

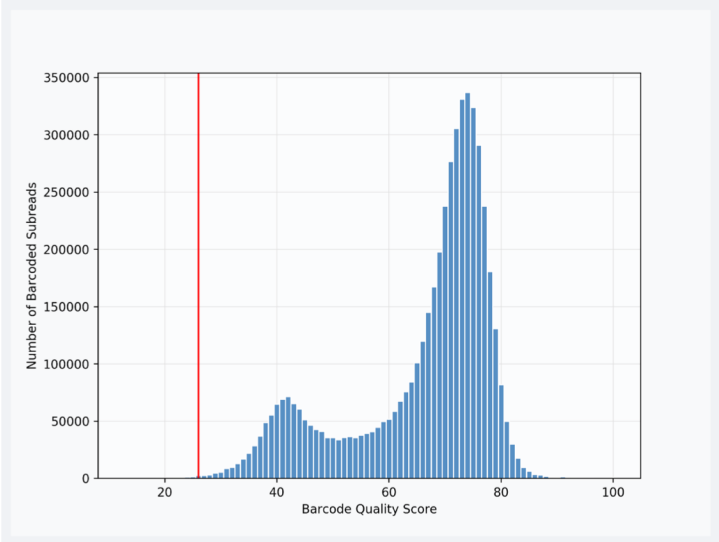
Analyze

### Barcode Data

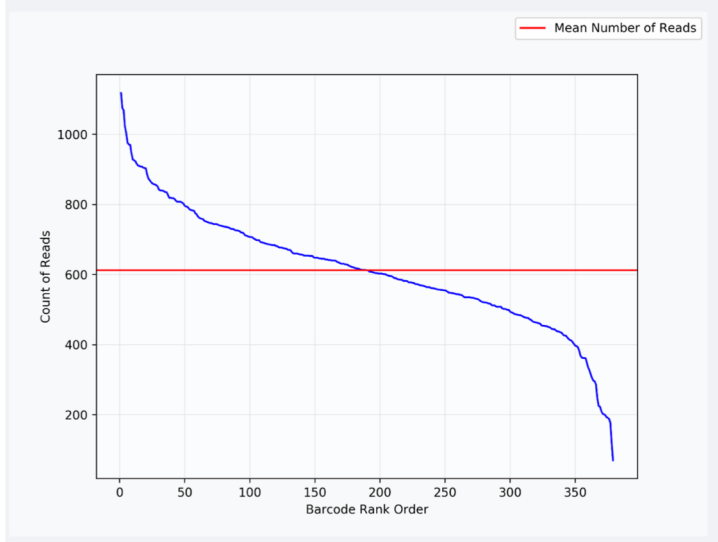
Bio Sample Name	Barcode Index	Barcode Name	Polymerase Reads	Subreads	Bases	Mean Read Length	Longest Subread Length	Mean Barcode Quality	Rank Order (Num. Reads)
Alice	0--0	lbc1--lbc1	2	9	17,251	10,659	2,958	55.0	2
Bob	1--1	lbc2--lbc2	2	5	6,878	6,414	2,433	55.0	1
Charles	2--2	lbc3--lbc3	1	10	20,529	26,336	2,858	59.0	3
No Name	None	Not Barcoded	1	2	3,146	3,181	2,737	0.0	NA

Barcode → Pool → Sequence → Demultiplex → Analyze

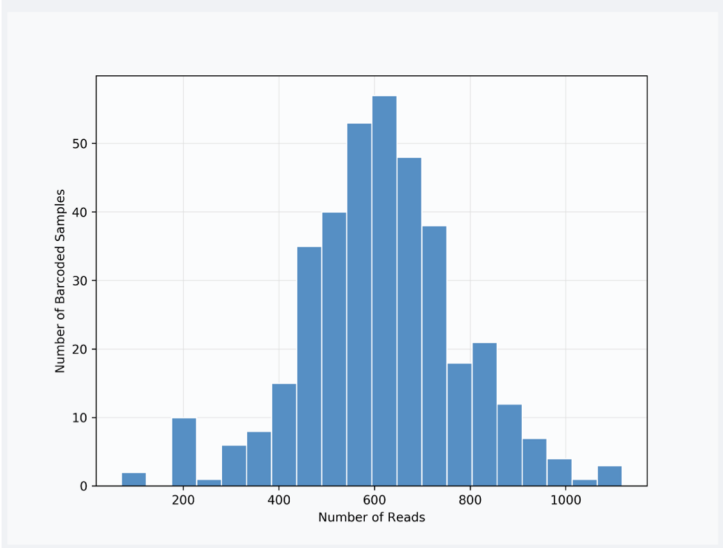
**Barcode Quality Score Distribution**



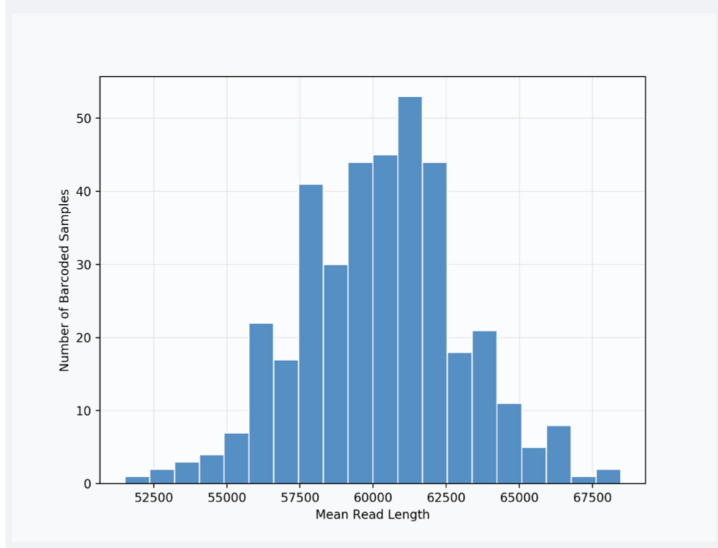
**Number Of Reads Per Barcode**



**Barcode Frequency Distribution**

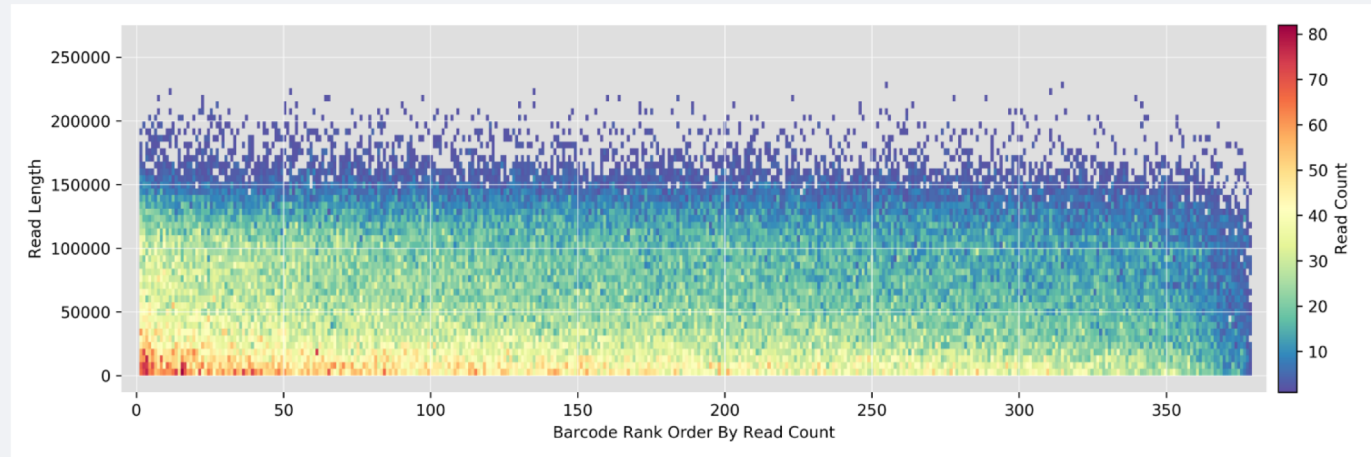


**Mean Read Length Distribution**





### Read Length Distribution By Barcode



### Barcode Quality Distribution By Barcode





### Create New Analysis - Settings

✕ CANCEL
▶ START AND CREATE NEXT

**Analysis Application \***  
 Minor Variants Analysis [Beta]

**Analysis Name \* ⓘ**  
 - KiwiTest01\_Cell1\_demux (lbc2--lbc2)

### Analysis of Multiple Datasets

View: BAM Data

**Analysis Type \***  
 One Analysis per Data Set - Custom Parameters

- One Analysis on All Data Sets
- One Analysis per Data Set - Identical Parameters
- One Analysis per Data Set - Custom Parameters

**Data Sets** Search... ☰

	Data Set Name	Well Sample Name	Bio Sample Name
▼	KiwiTest01_Cell1_demux <span>All</span> <span>None</span>	96-01-1L-10pM	96-01-1L-10pM
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc1--lbc1)	96-01-1L-10pM	lbc1--lbc1
<input checked="" type="checkbox"/>	KiwiTest01_Cell1_demux (lbc2--lbc2)	96-01-1L-10pM	lbc2--lbc2
<input checked="" type="checkbox"/>	KiwiTest01_Cell1_demux (lbc3--lbc3)	96-01-1L-10pM	lbc3--lbc3
<input checked="" type="checkbox"/>	KiwiTest01_Cell1_demux (lbc4--lbc4)	96-01-1L-10pM	lbc4--lbc4
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc5--lbc5)	96-01-1L-10pM	lbc5--lbc5
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc6--lbc6)	96-01-1L-10pM	lbc6--lbc6
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc7--lbc7)	96-01-1L-10pM	lbc7--lbc7
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc8--lbc8)	96-01-1L-10pM	lbc8--lbc8
<input type="checkbox"/>	KiwiTest01_Cell1_demux (lbc9--lbc9)	96-01-1L-10pM	lbc9--lbc9

### Minor Variants

**Target Config ⓘ**  
 [Empty field]

**ADVANCED ANALYSIS PARAMETERS**

# DEMULTIPLEXING POWERED BY LIMA

## Features:

- Process both, raw subreads and CCS reads
- PacBio BAM in- and output (optional: split output BAM files by barcode)
- Extensive reports that allow in-depth quality control
- Clip barcode sequences and annotate bq and bc tags
- Agnostic of input barcode sequence orientation
- Infer subset of barcodes in sample
- Latest develop version as binary

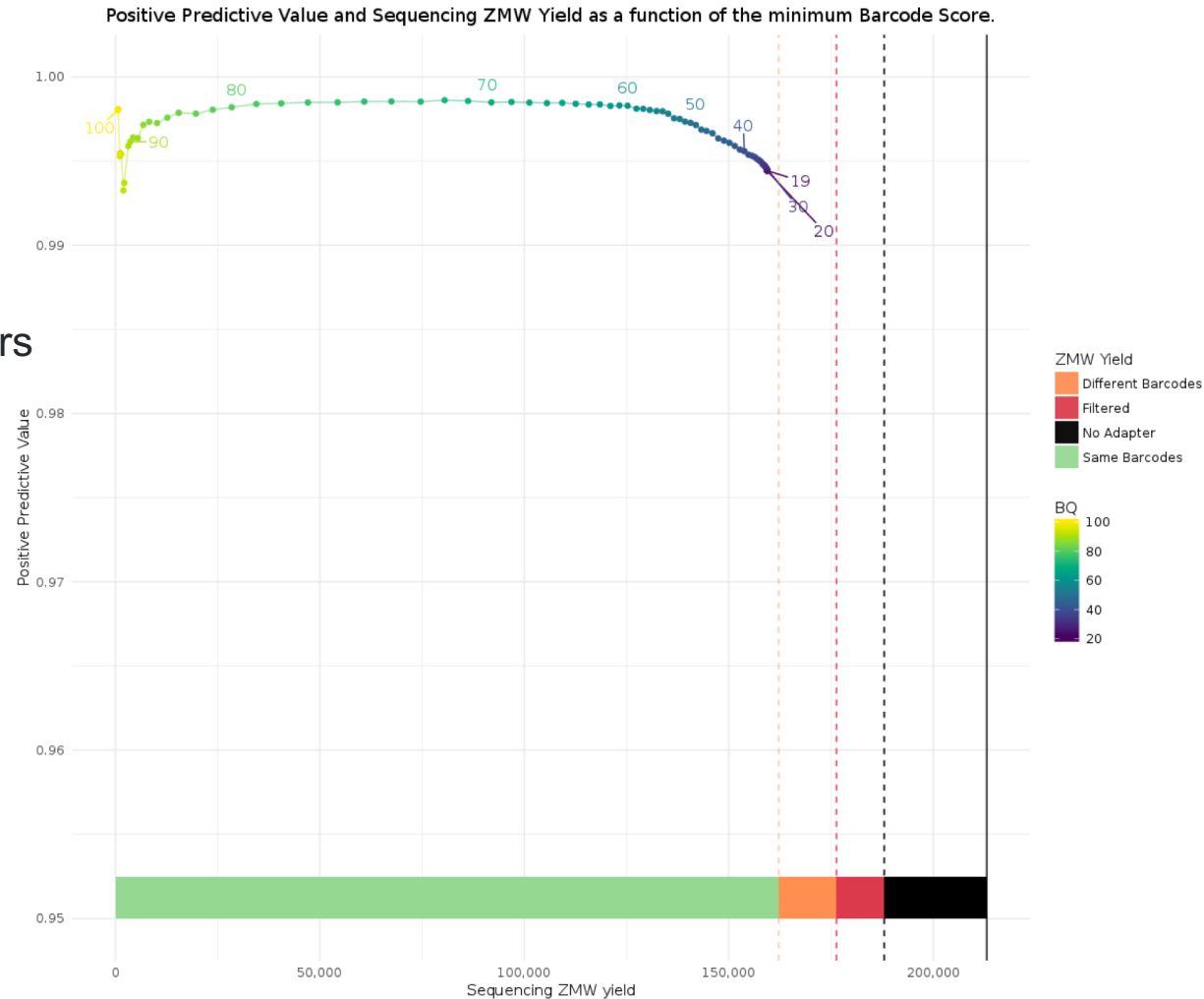
```
conda install -c bioconda lima
```

- Public documentation

```
https://github.com/PacificBiosciences/barcoding
```

# DEMULTIPLEXING POWERED BY LIMA - PERFORMANCE

- 2 kb amplicons
- Vector-sequence-specific primers
- 6 hours long movies
- 2 hours pre-extension
- 384 barcodes
- symmetric
  
- Instrument version is 5.0.0  
and the chemistry is S/P2-C2



## DEMULTIPLICING POWERED BY LIMA - PERFORMANCE

- **Symmetric has best performance**
- **Asymmetric for short amplicons**
- **Do not mix designs  
(not supported in SMRT Link)**

Design	Plex	PPV %
Symmetric	8	99.7
	16	99.7
	40	99.6
	48	99.5
	96	99.4
	384	99.1
Asymmetric	28	98.8
	384	97.0
Sym+Asym	36	90.6



[www.pacb.com](http://www.pacb.com)

For Research Use Only. Not for use in diagnostics procedures. © Copyright 2018 by Pacific Biosciences of California, Inc. All rights reserved. Pacific Biosciences, the Pacific Biosciences logo, PacBio, SMRT, SMRTbell, Iso-Seq, and Sequel are trademarks of Pacific Biosciences. BluePippin and SageELF are trademarks of Sage Science. NGS-go and NGSengine are trademarks of GenDx. FEMTO Pulse and Fragment Analyzer are trademarks of Advanced Analytical Technologies.

All other trademarks are the sole property of their respective owners.