# Targeting

**Chris Mattocks**
**National Genetics Reference Laboratory (Wessex)**


# Diagnostic implementation

- Targeting specific regions of interest
  - PCR
  - Arrays
  - Probe based methods
- Analysing multiple patient samples in parallel
  - Physical separation
  - ID tagging
  - Pooling strategies
- Laboratory process
  - Rationalising preparation / amplification process
  - Turn around times
  - Utilisation of capacity
- Data handling and analysis pipeline

# Potential diagnostic approaches

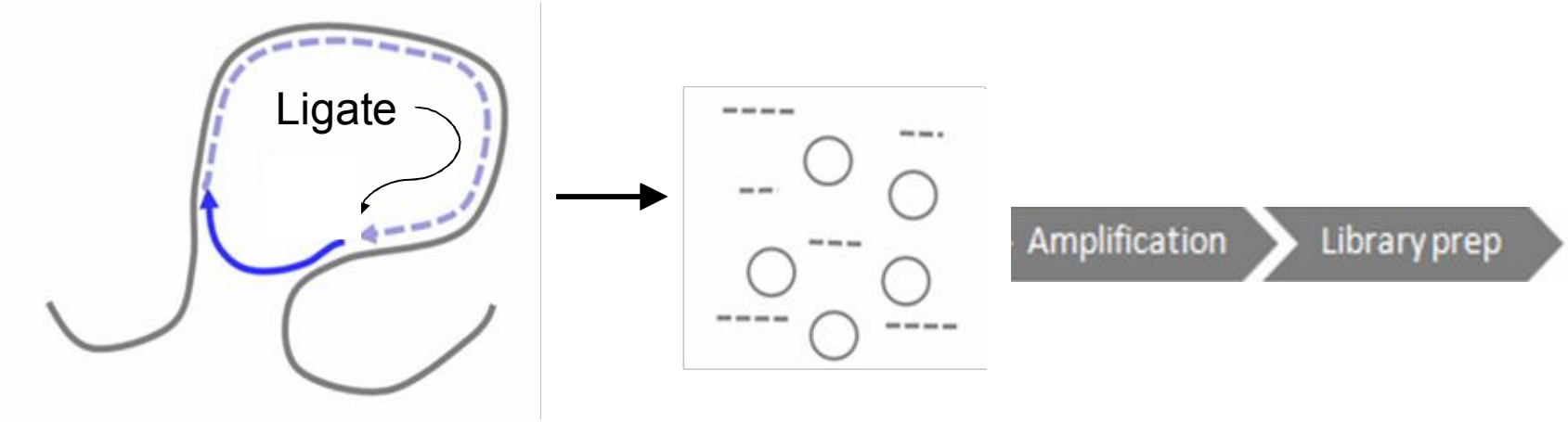
| Type of partitioning | Sequence | Example | For | Against |
|---|---|---|---|---|
| Whole genome sequencing (none) | Entire genome | PacBio<br>Complete genomics<br>Oxford Nanopore<br>?? | •Simplicity<br>•Cost?<br>•Turnaround?<br>•Pre-emptive | •Current feasability<br>•Colateral findings<br>•Data analysis<br>•Cost? |
| Standardised genomic partitioning | Fixed set of genes | Nimblegen | •Simplicity<br>•Availability<br>•Pre-emptive? | •Flexibility<br>•Collateral findings<br>•Data analysis |
| Flexible genomic partitioning | Selection based on referral | Probe based targeting | •Flexibility<br>•Patient based<br>•Efficiency | •Development<br>•Sample processing |

# Standardised partitioning

- Solid phase array:
  - Current Nimblegen array 5Mb
  - >1000 genes
  - Cost per requirement

- Liquid phase array:
  - Probe based approach
  - Probes generated on an array

# Highly multiplexed amplification



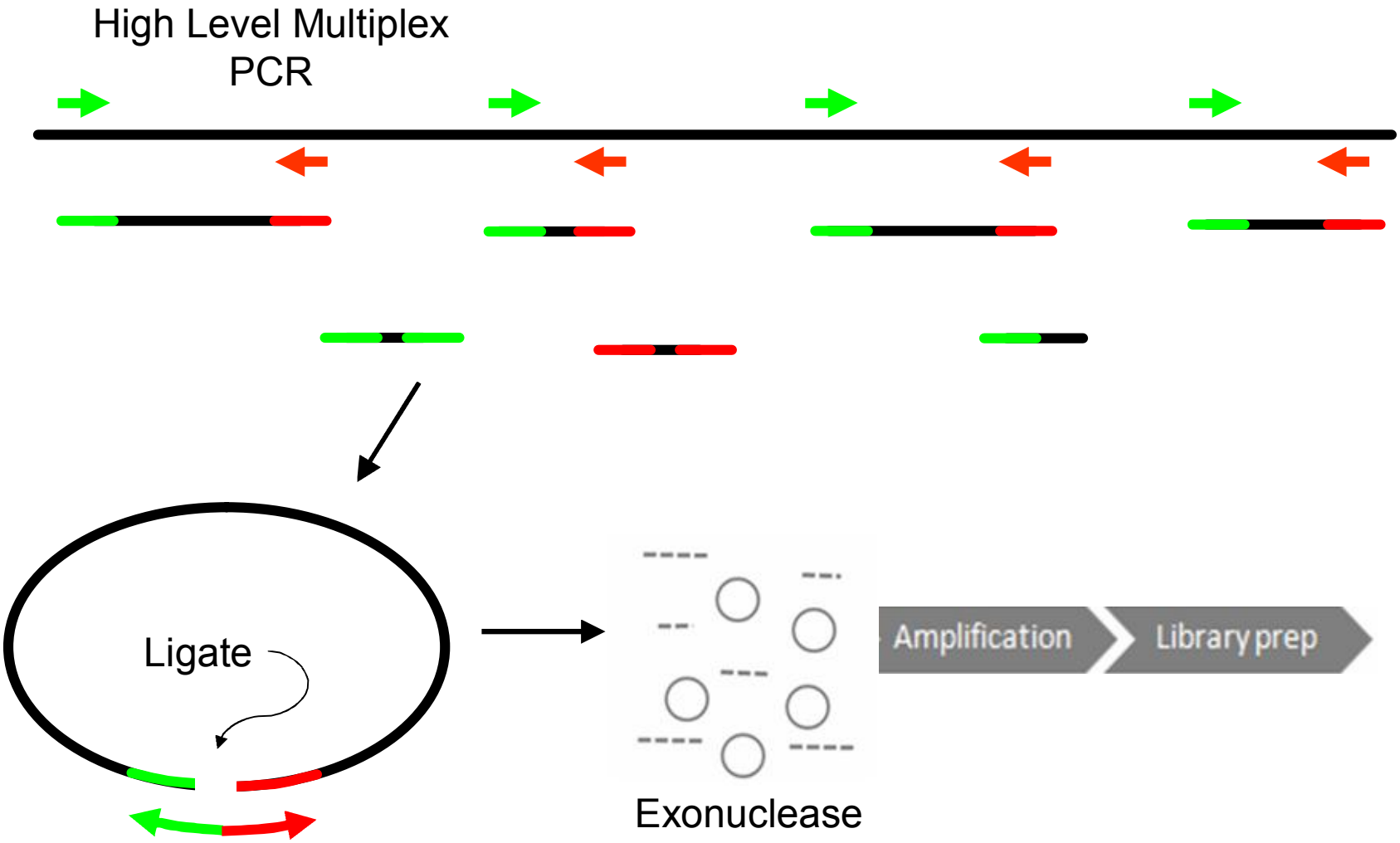

**Multiplex amplification of large sets of human exons**
Gregory J Porreca, Kun Zhang, Jin Billy Li, Bin Xie, Derek Austin, Sara L Vassallo, Emily M LeProust, Bill J Peck, Christopher J Emig, Fredrik Dahl, Yuan Gao, George M Church & Jay Shendure
*Nat Methods. 2007 Nov;4(11):931-6*

# Flexible partitioning

- Even if whole genome is desirable only ~2% of genome is coding
- Collateral findings
- Reporting issues – current understanding
- Greater capacity (e.g DNA vs RNA)
- Cost

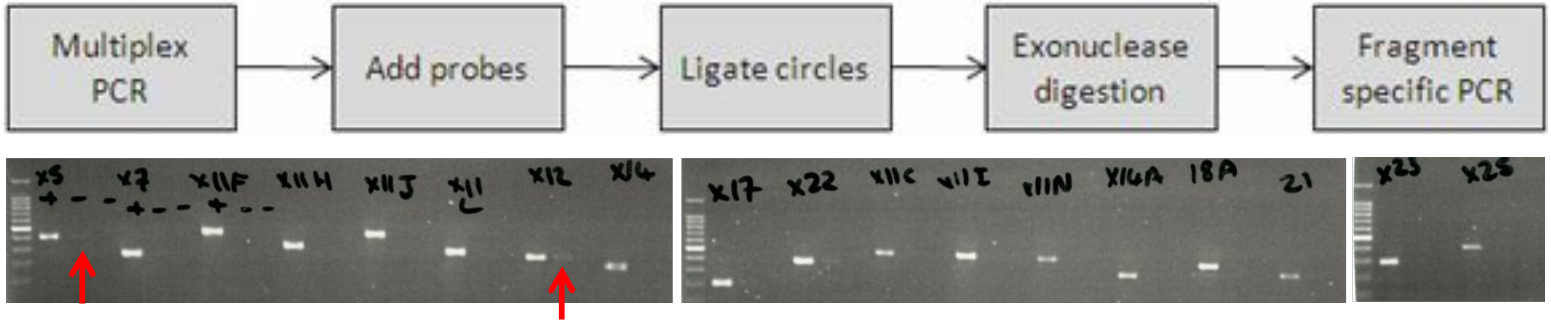
£ partitioning process < £ genome sequence

# Gene-Collector




**Multiplex amplification of all coding sequences within 10 cancer genes by Gene-Collector.**
Simon Fredriksson,* Johan Banér, Fredrik Dahl, Angela Chu, Hanlee Ji, Katrina Welch, and Ronald W. Davis
*Nucleic Acids Res.* 2007 April; 35(7): e47

# 18 plex circularisation (BRCA1 &BRCA2)





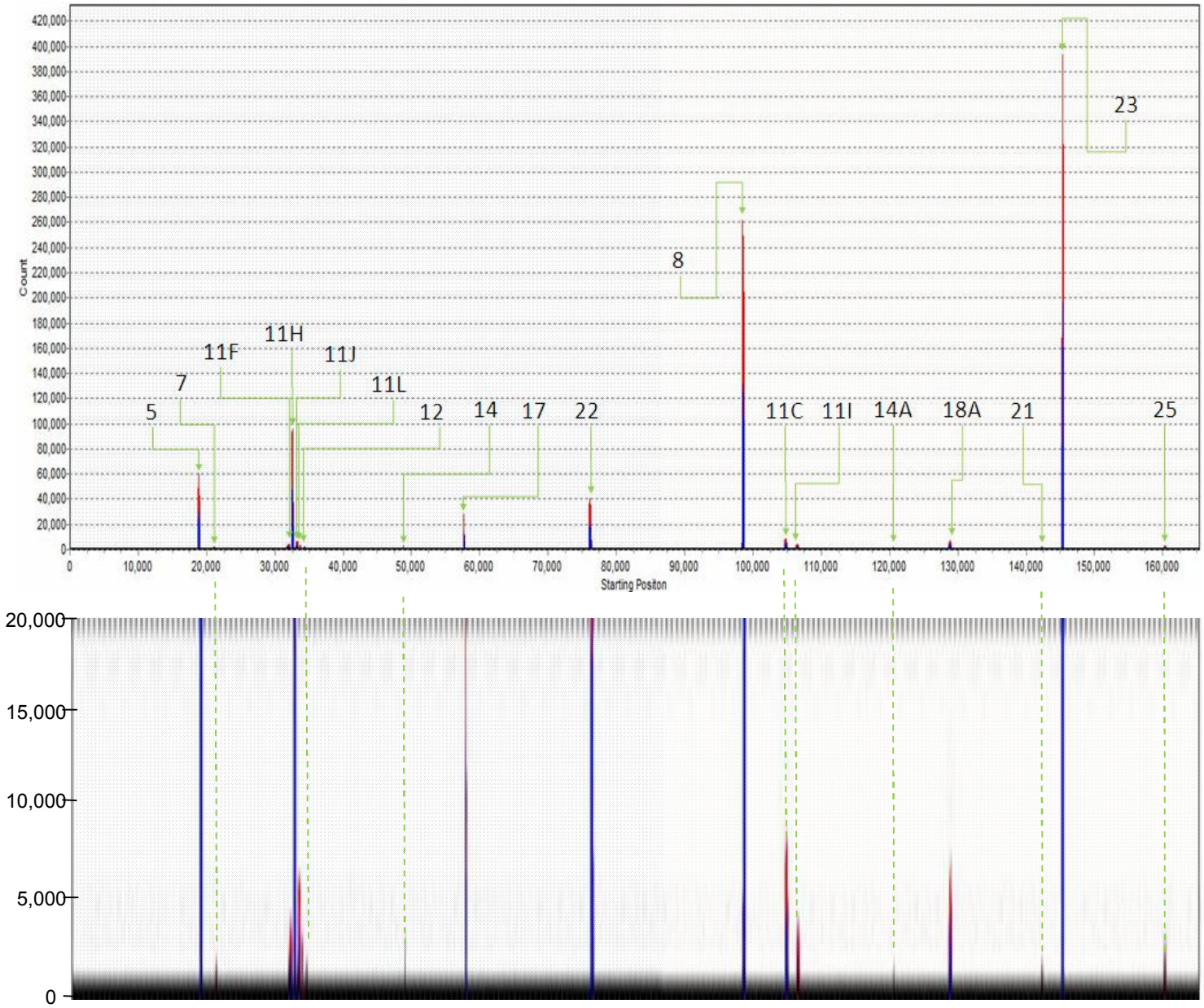

| | Multiplex primers | Circularisation probes | BC1-05 | BC1-07 | BC1-11F | BC1-11H | BC111J | BC1-11L | BC1-12 | BC1-14 | BC1-17 | BC1-22 | BC2-08 | BC2-11C | BC2-11I | BC2-14A | BC2-18A | BC2-21 | BC2-23 | BC2-25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test | + | + | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ | +++ |
| Multiplex control | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| Circularisation control | + | - | + | - | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - |

- Successfully recovered all 18 specific
- 2 /18 showed low level recovery from reaction with no circularisation probe indicating digestion of non-specific [linear] products was not complete

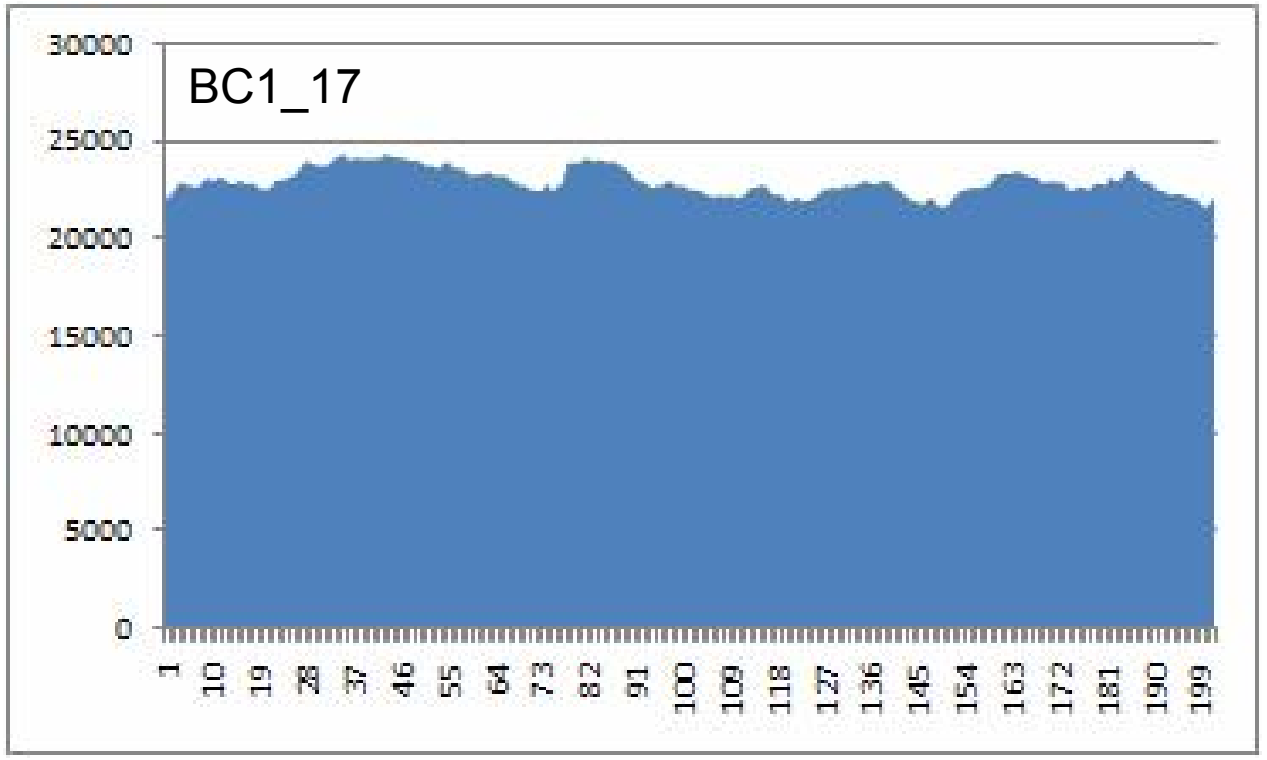
# 18 plex circularisation – Illumina run

Representation

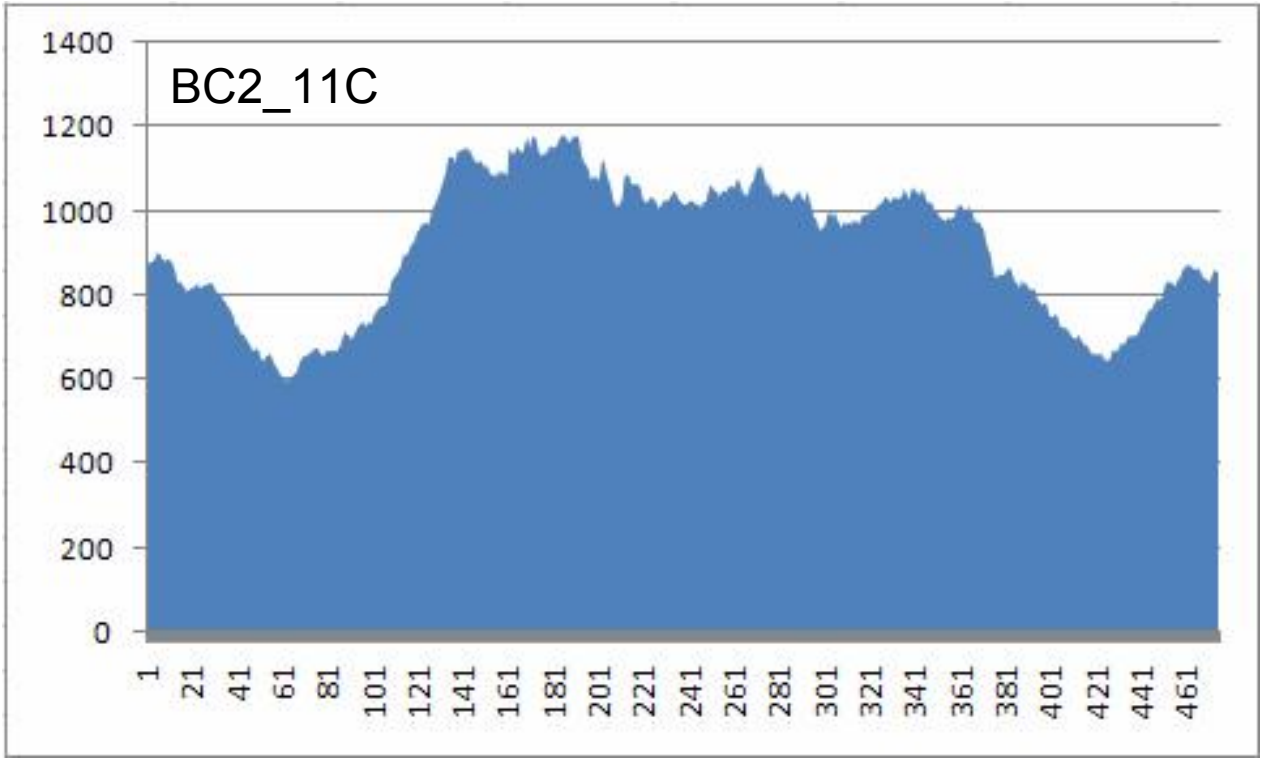
| Gene | Exon | Multiple of minimum |
|---|---|---|
| 1 | 11H | 52 |
| 1 | 5 | 32 |
| 1 | 22 | 21 |
| 1 | 17 | 16 |
| 1 | 11J | 4 |
| 1 | 11F | 2 |
| 1 | 11L | 2 |
| 1 | 14 | 2 |
| 1 | 7 | 1 |
| 1 | 12 | 1 |
| 2 | 23 | 208 |
| 2 | 8 | 14 |
| 2 | 11C | 5 |
| 2 | 18A | 4 |
| 2 | 11I | 2 |
| 2 | 25 | 2 |
| 2 | 14A | 1 |
| 2 | 21 | 1 |



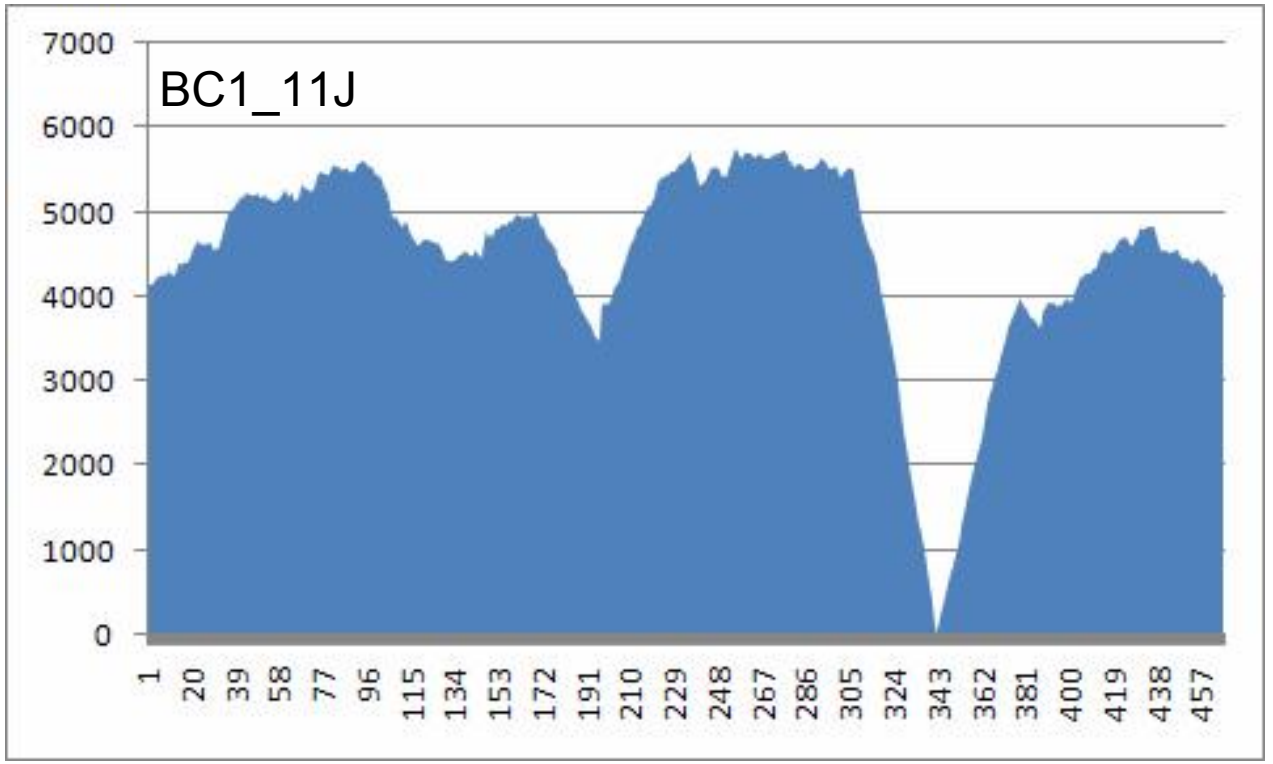
# Example data



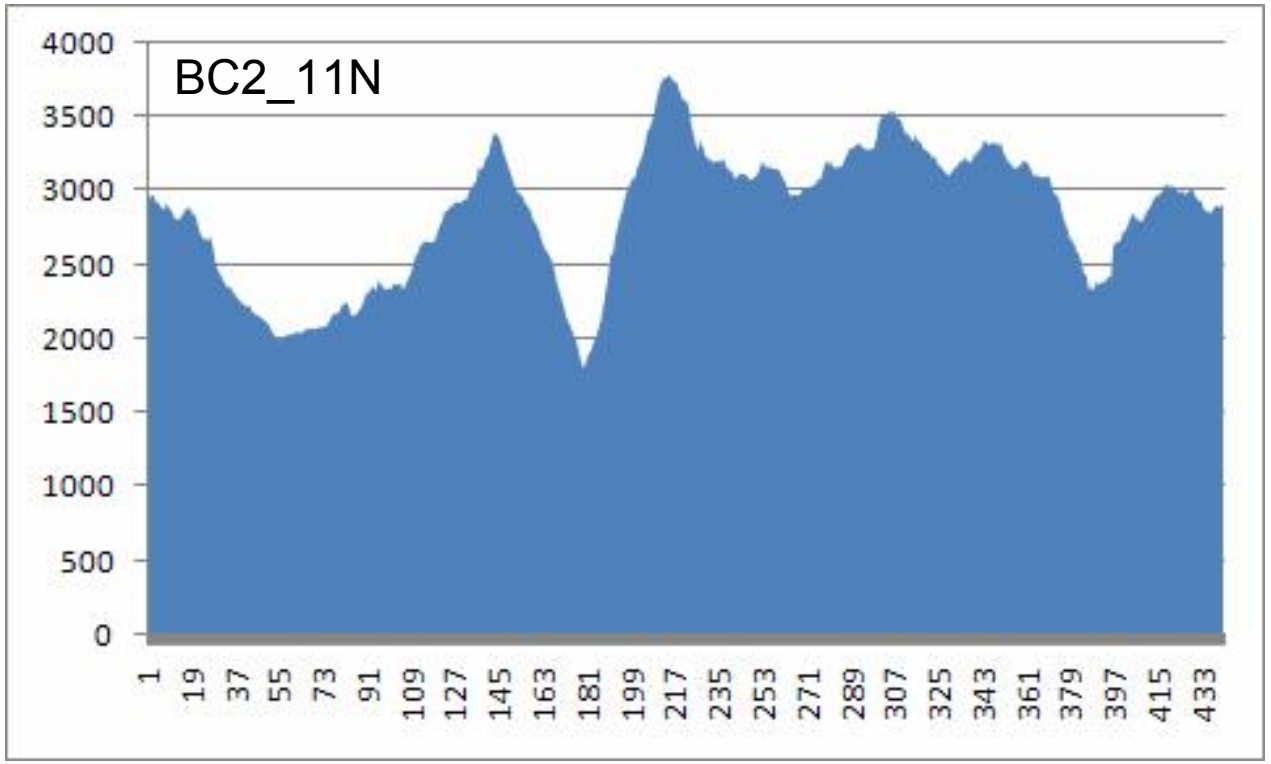
# Example data



# Example data

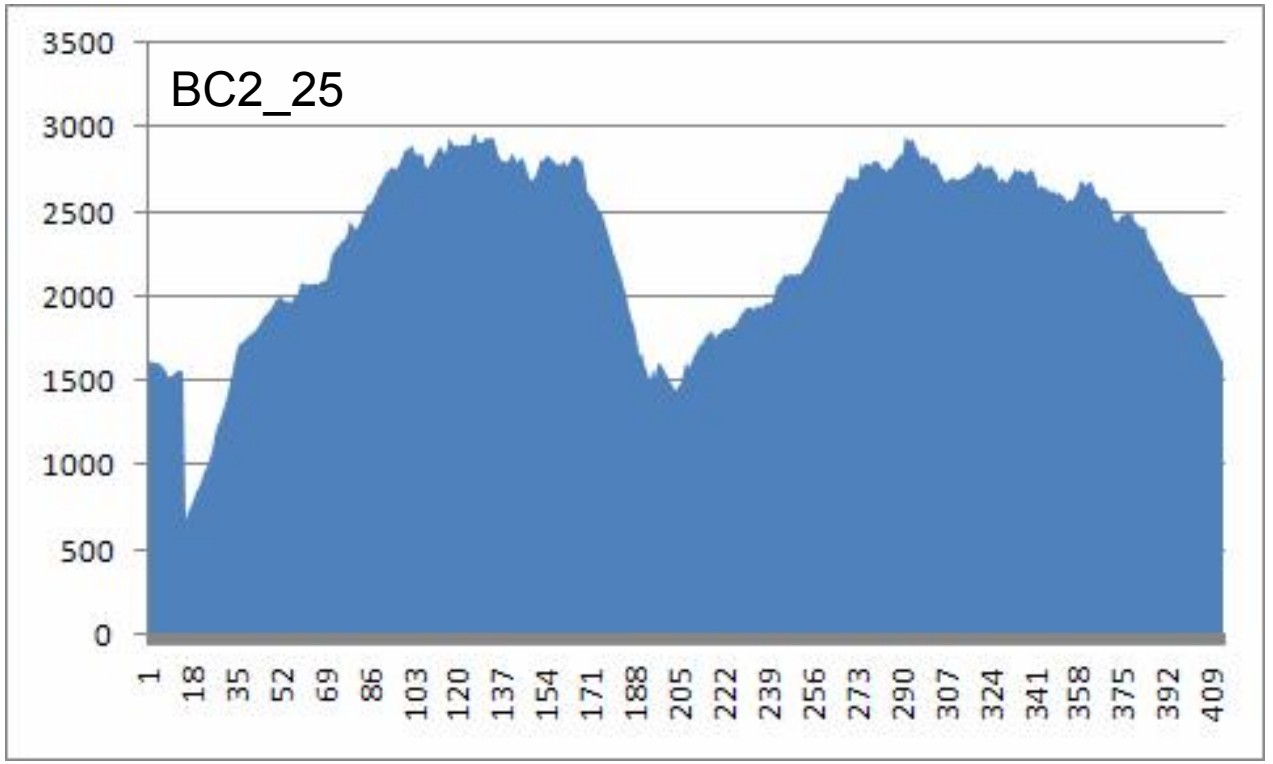


# Example data



# Example data

# Example data



# Targeting by circularisation

**Progress**

- Initial 18 plex circularisation promising
- Currently ~70% of sequence maps to ROI – unmappable sequence thought to arise from use of random primers in post circularisation amplification.
  a) Looking at use of universal specific primers for post circularisation amplification
  b) More stringent digestion regime

- Currently ~200 fold difference representation between highest and lowest – differential though to arise from amplification stages.
  a) Minimise multiplex PCR cycling to minimise amplification biases
  b) Maximise multiplex PCR cycling and use circularisation probe concentration to normalise target representation
  c) Real time assays to assess representation

# Next steps

- Investigate effect of fragment length
- Investigate methods for introducing ID tags
- Ramp up to full BRCA screen

# NGRL work plan


- Design phase: Compile a list of clinically interesting genes (~100)
  - Including reference genes currently tested (BRCA1, BRCA2, FBN1)
  - Gene pathways
- Pilot phase: Develop and optimise circularisation assays using BRCA genes as model
  - Normalisation
  - Coverage
  - Sequence quality
- Testing phase: Prospective study using array and probe approaches sample run on all instruments and compared to diagnostic result.
  - Expanded probe set