



# 454 Sequencing in Molecular Diagnostics

---

**Ann Curtis and Jonathan Coxhead**

NewGene Ltd

In partnership with Newcastle Hospitals and the  
University of Newcastle



# Next Generation Sequencing in Molecular Genetics

---

- Workflow
- Gene/disorders we have worked on
- Data
- Problems



# Next Generation Sequencing in Molecular Genetics

---

## Move from

- Sanger chain termination sequencing
  - exon by exon
  - most likely candidate genes
  - individual patients

## To

- Parallel sequencing (various chemistries)
  - all exons
  - many genes
  - many patients



# 454 Sequencing in Molecular Genetics

---

Two approaches to clinical sequencing–

- Amplicon sequencing
  - sequencing of PCR products
  - extension of current Sanger methods but much higher through put
- Sequence capture
  - sequencing of genomic DNA captured onto a custom designed chip

# 454 GS-FLX



## Capacity – Titanium chemistry (pyrosequencing)

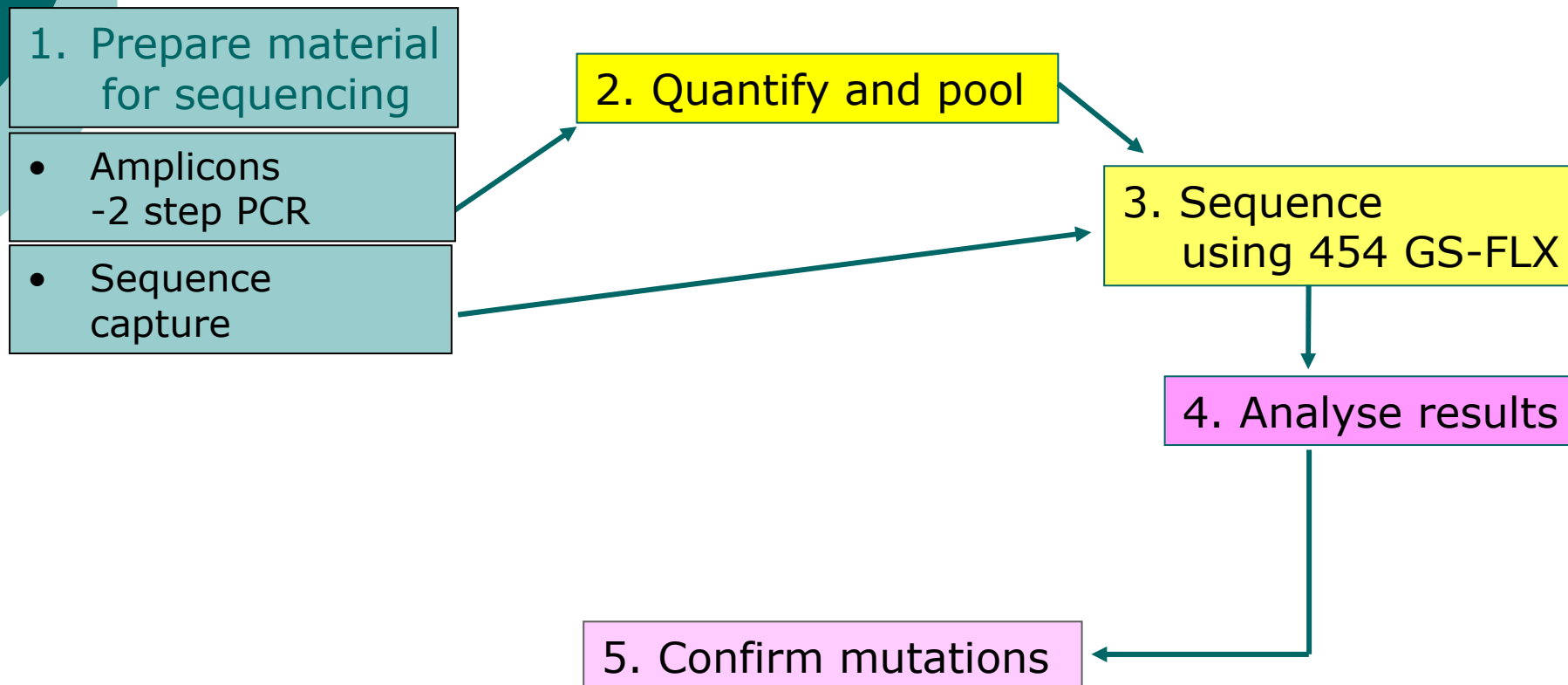
400 – 600 Mbases / run

400 – 600 bp PCR products (amplicons)

~1 million reads = 20,000 amplicons @ 50x coverage

10 hrs

# Clinical sequencing workflow





# Why use amplicon sequencing

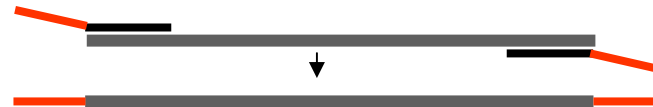
---

- Based on PCR – familiar technique
- Good use of resources - staff experience - equipment already available
- Minimal capital investment
- Test experimental design
- Common workflow for all genes
- Flexibility
- Cost efficient
- Scalable up to a point

Workflow section 1

# Two-step PCR for flexibility

- PCR STEP 1- Can perform PCR in multiplex
  - all primers have gene specific sequence and **universal (M13) tag**



- PCR STEP 2-
  - need to add patient specific 'bar code' (MID) and **Roche sequencing tag (A and B adaptors)**



- Require sufficient MID primers to allow appropriate mixing of patient samples



# Workflow section 1

## Two-step PCR using robotics

1	9	17	25	33	41	1	9	17	25	33	41
2	10	18	26	34	42	2	10	18	26	34	42
3	11	19	27	35	43	3	11	19	27	35	43
4	12	20	28	36	44	4	12	20	28	36	44
5	13	21	29	37	45	5	13	21	29	37	45
6	14	22	30	38	46	6	14	22	30	38	46
7	15	23	31	39		7	15	23	31	39	
8	16	24	32	40		8	16	24	32	40	

Primer plate  
Primers in multiplex groups  
Prepared, checked, stored

transfer

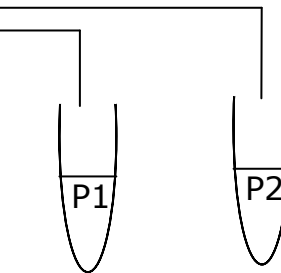
1	9	17	25	33	41	1	9	17	25	33	41
2	10	18	26	34	42	2	10	18	26	34	42
3	11	19	27	35	43	3	11	19	27	35	43
4	12	20	28	36	44	4	12	20	28	36	44
5	13	21	29	37	45	5	13	21	29	37	45
6	14	22	30	38	46	6	14	22	30	38	46
7	15	23	31	39		7	15	23	31	39	
8	16	24	32	40		8	16	24	32	40	

1° PCR plate

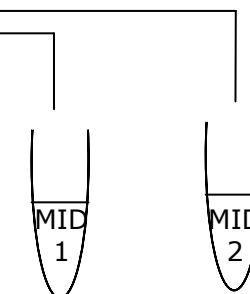
transfer

1	9	17	25	33	41	1	9	17	25	33	41
2	10	18	26	34	42	2	10	18	26	34	42
3	11	19	27	35	43	3	11	19	27	35	43
4	12	20	28	36	44	4	12	20	28	36	44
5	13	21	29	37	45	5	13	21	29	37	45
6	14	22	30	38	46	6	14	22	30	38	46
7	15	23	31	39		7	15	23	31	39	
8	16	24	32	40		8	16	24	32	40	

2° PCR plate

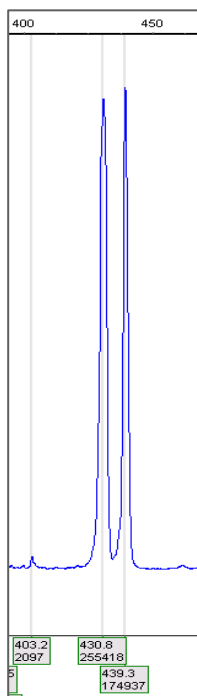


Master mixes with patient DNA

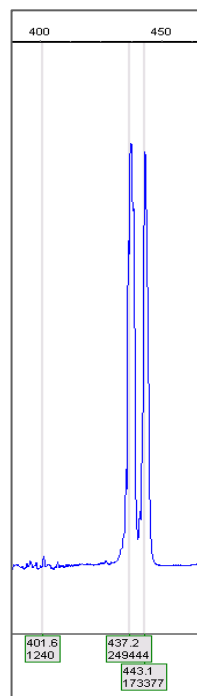


Master mixes with MID primers

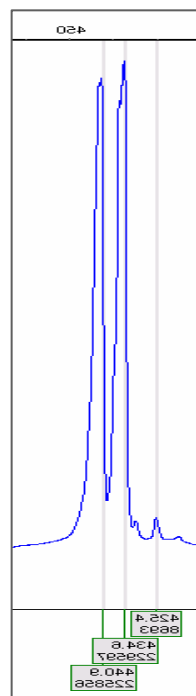
# Diplex PCR for amplicon sequencing



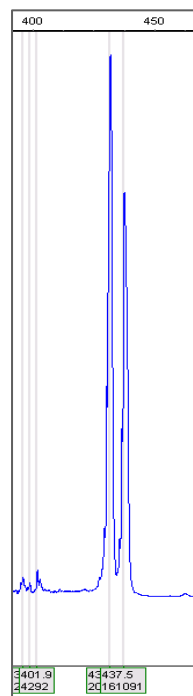
VCP ex2  
SGCB ex3



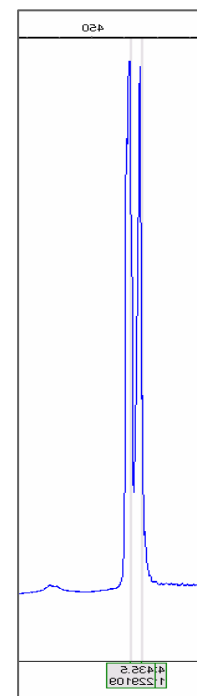
CAPN ex13  
SGCB ex6



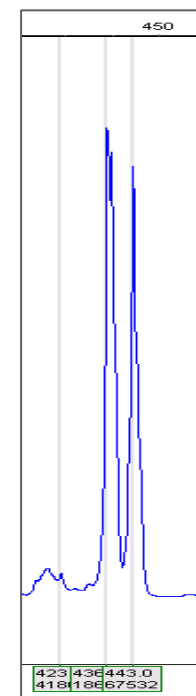
DYSF ex45  
SGCG ex8



MYOT ex7  
CRYAB ex 3



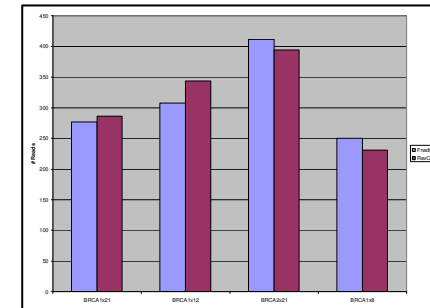
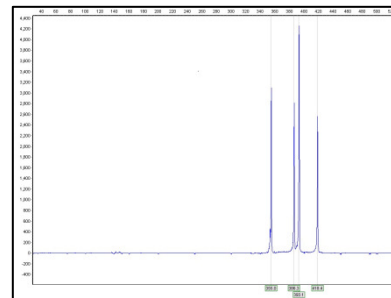
CRYAB ex1  
SGCG ex2



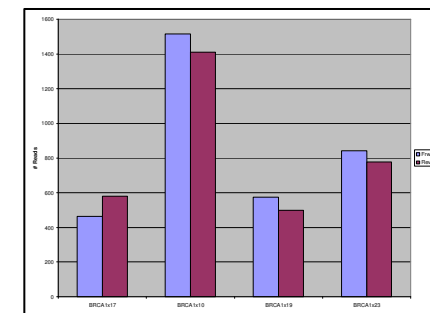
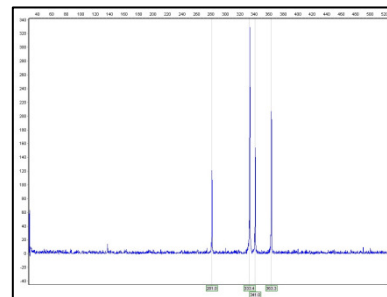
FLNC ex41  
FHL1 ex6

# Multiplex two step PCR

BRCA1 ex21,  
BRCA1 ex12,  
BRCA2 ex21,  
BRCA1 ex8



BRCA1 ex17,  
BRCA1 ex10,  
BRCA2 ex19,  
BRCA1 ex23



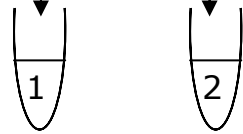
Multiplex reactions  
on ABI

Sequencing reads  
from GS-FLX

# Workflow section 2: Quantify and pool PCR products

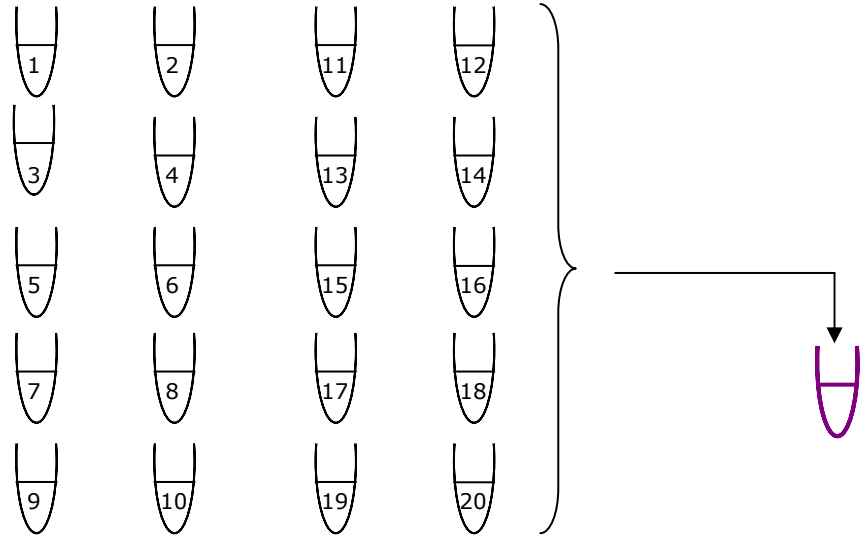
2° PCR plate

1	9	1	7	2	5	3	3	4	1	1	9	1	7	2	5	3	3	4	1
2	1	0	1	8	2	6	3	2	2	2	1	0	1	8	2	6	3	2	2
3	1	1	1	9	3	5	4	3	3	1	1	1	9	3	5	4	3	3	1
4	1	3	7	0	8	3	6	4	4	4	1	2	2	2	3	5	6	4	4
5	3	3	1	2	9	3	7	4	5	5	1	3	0	2	9	3	7	4	5
6	2	4	2	5	0	3	8	4	6	4	2	2	3	0	3	8	4	6	4
7	1	5	2	3	3	1	3	9	7	1	1	5	2	3	3	1	3	9	7
8	1	6	2	4	3	2	4	0	8	1	1	6	2	4	3	2	4	0	8



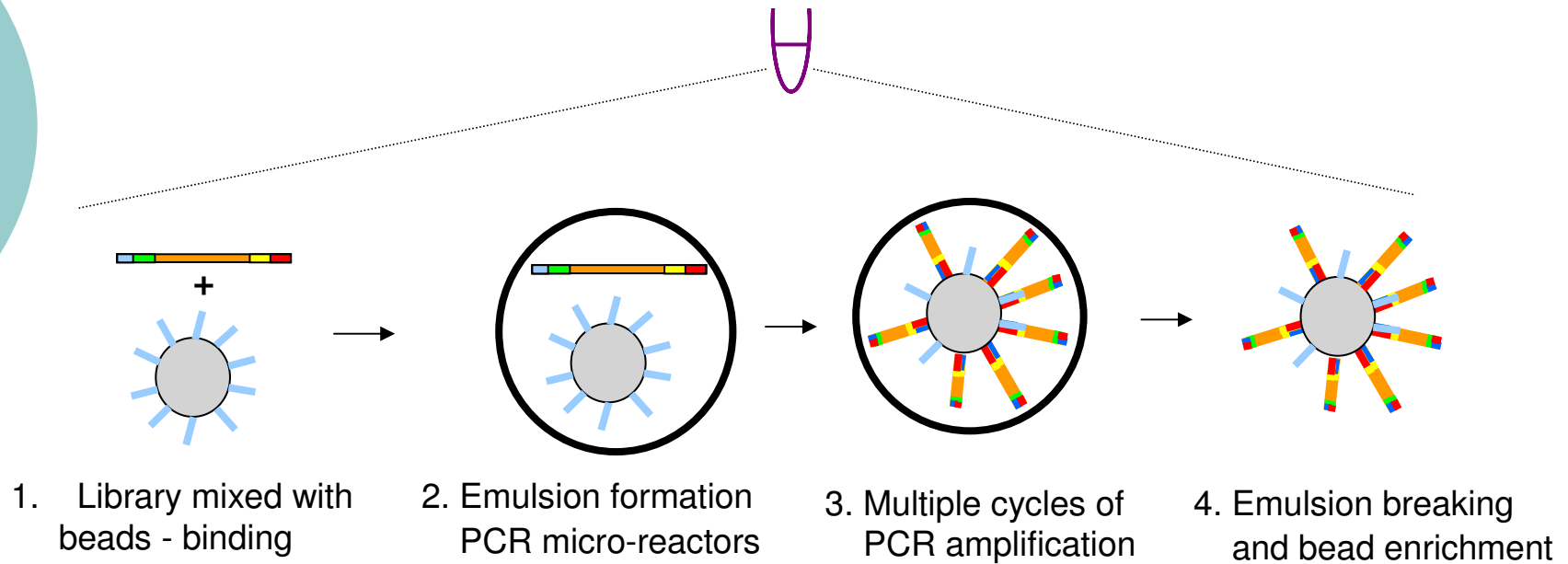
Quantify using pico green and fluorimetry

Pool PCRs for each patient into 'library'

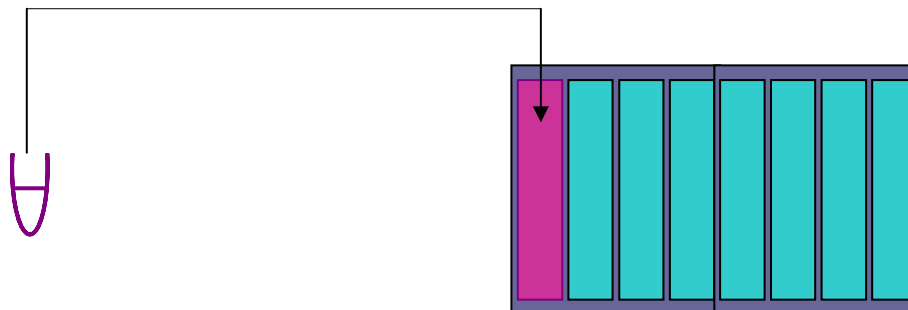


Pool many patient libraries into a single library

# Clonal amplification – emulsion PCR

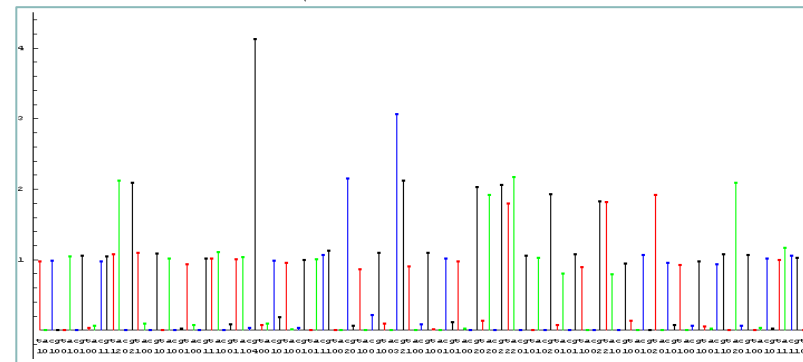
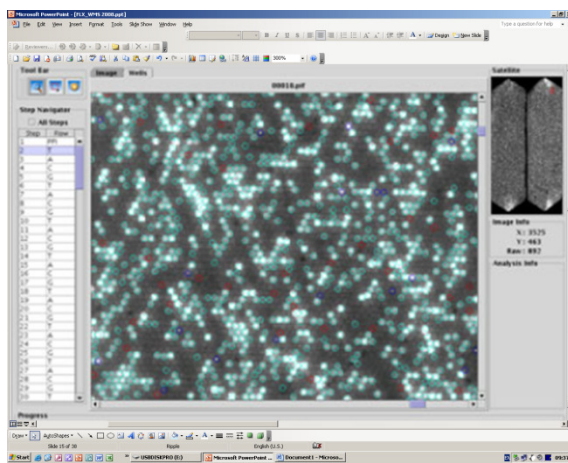
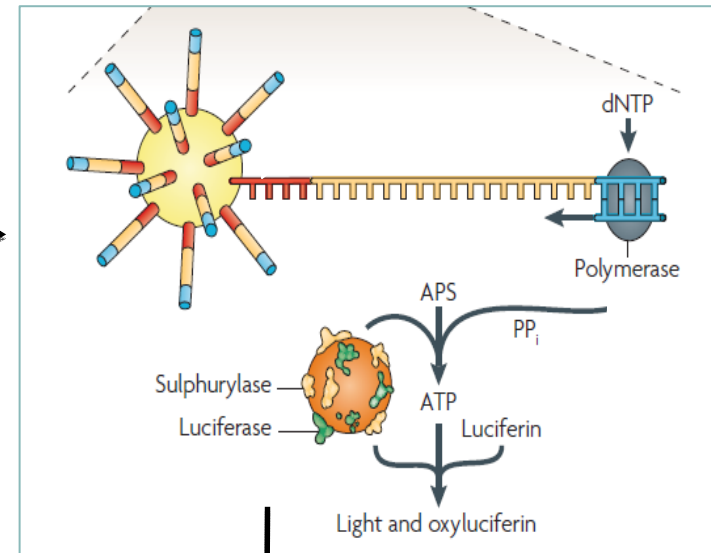
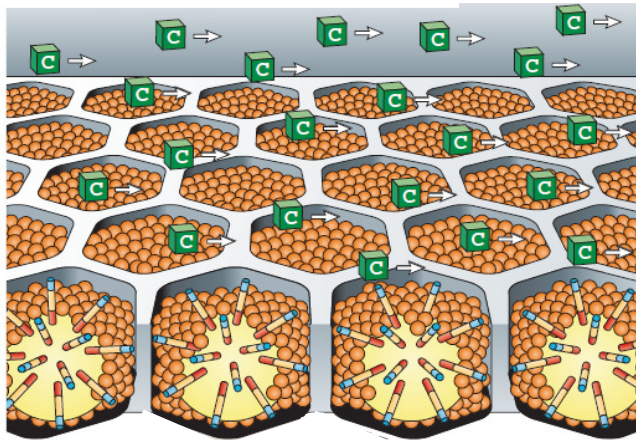


Load into 1/8<sup>th</sup> of FLX plate



2000 PCR products  
20 patients  
BRCA1 + BRCA2

# Sequencing





# Data

---

## **Analysis software**

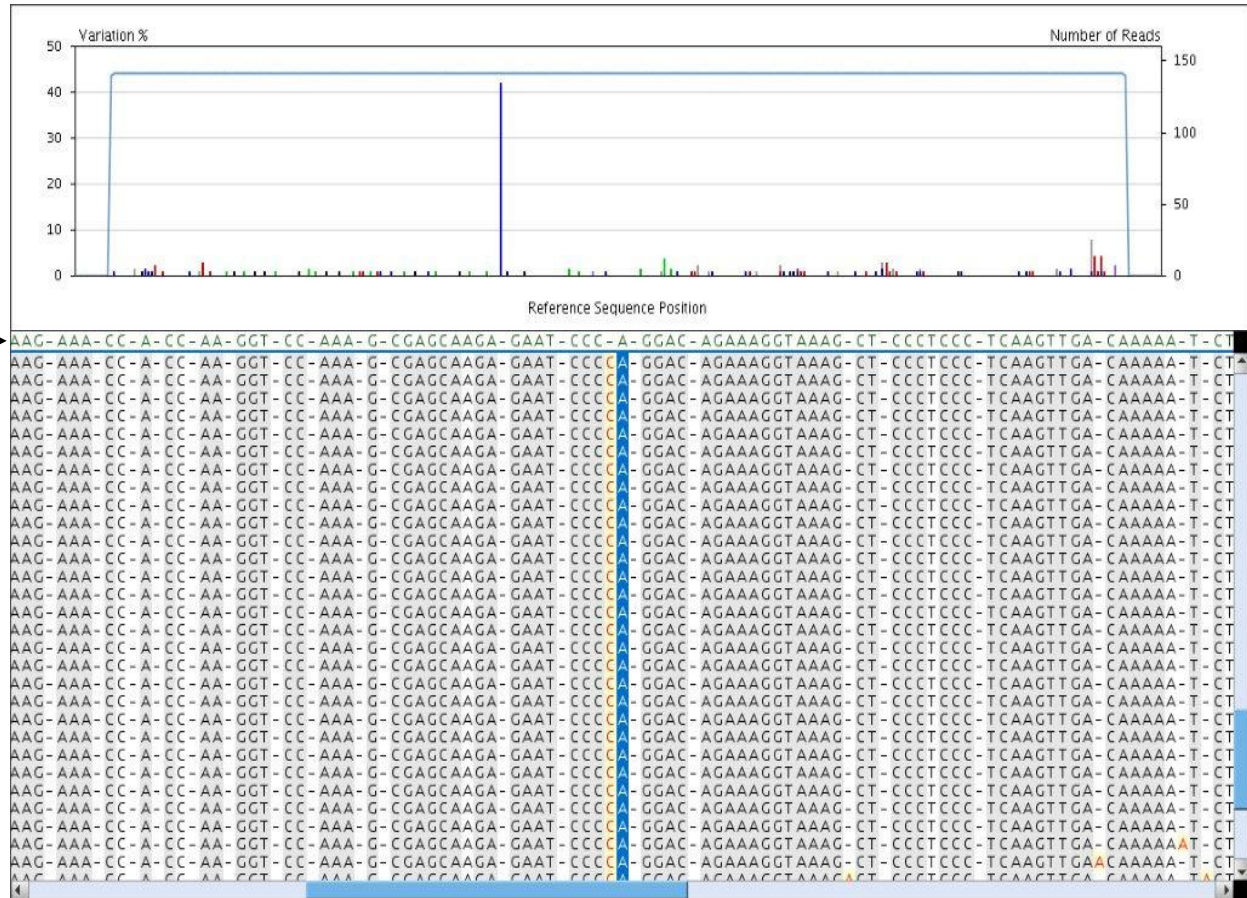
- GS De Novo Assembler
- GS Reference Mapper
- GS Amplicon Variant Analyzer

# Data output

## Alignment of individual reads

Reads aligned to reference sequence

- Software searches for variants and highlights in variant plot

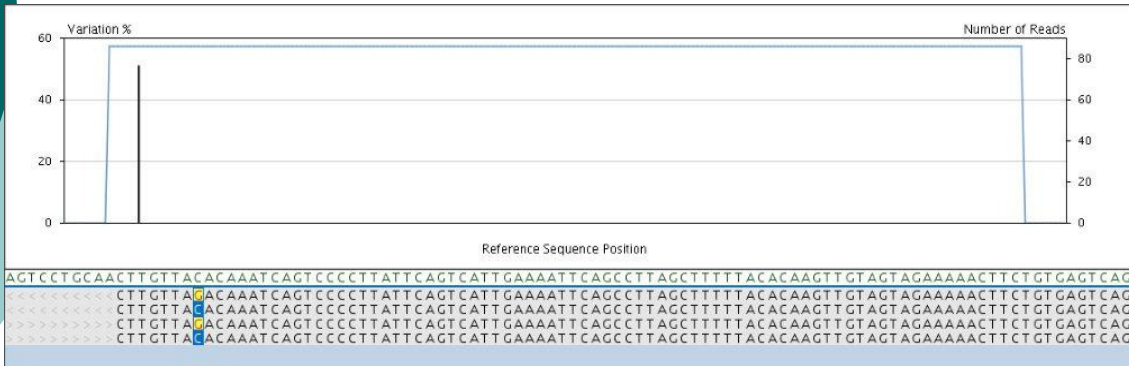




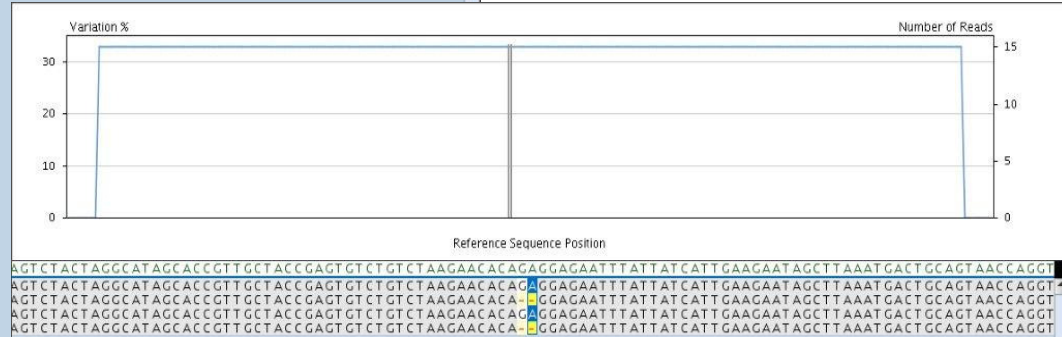
# Data output – variant table

Reference	Variant	Max	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...	Sample_M...
Exon_x2	48:T/C	52.38	0.00	-	52.38	-	▼ 50.00	0.00	-	▼ 0.00	-	-
Exon_x3	44:A/T	▼ 50.00	7.50	0.00	0.00	▼ 50.00	▼ 50.00	0.00	0.00	0.00	0.00	26.09
Exon_x3	46:C/A	▼ 50.52	7.50	0.00	0.00	▼ 50.52	▼ 50.00	0.00	0.00	0.00	0.00	26.09
Exon_x3	47:T/C	▼ 50.00	7.50	0.00	0.00	▼ 49.48	▼ 50.00	0.00	0.00	0.00	0.00	26.09
Exon_x3	161:G/A	52.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	52.17
Exon_x5	62:A/T	▼ 77.59	▼ 69.41	▼ 77.59	▼ 62.96	4.76	0.00	▼ 75.00	▼ 73.68	▼ 52.17	▼ 25.00	-
Exon_x5	65:T/A	▼ 77.59	▼ 69.41	▼ 77.59	▼ 61.48	4.76	0.00	▼ 75.00	▼ 73.68	▼ 52.17	▼ 25.00	-
Exon_x5	67:A/T	▼ 55.56	0.00	0.00	0.00	▼ 55.56	▼ 50.00	0.00	0.00	0.00	0.00	-
Exon_x5	70:T/A	▼ 55.56	0.00	0.00	0.00	▼ 55.56	▼ 50.00	0.00	0.00	0.00	0.00	-
Exon_x8	100:C/T	40.48	0.00	0.00	0.00	40.48	▼ 0.00	0.00	-	0.00	0.00	-
Exon_x8	251:A/T	22.22	0.00	0.00	0.00	3.57	▼ 0.00	0.00	-	0.00	22.22	-
Exon_x8	252:G/A	22.22	0.00	0.00	0.00	3.57	▼ 0.00	0.00	-	0.00	22.22	-
Exon_x8	254:T/G	22.22	0.00	0.00	0.00	3.57	▼ 0.00	0.00	-	0.00	22.22	-
Exon_x11_5	200:A/T	▼ 49.48	▼ 18.92	▼ 15.45	▼ 21.52	▼ 20.75	▼ 30.00	▼ 21.74	0.00	▼ 26.10	▼ 49.48	-
Exon_x11_5	201:T/A	▼ 49.48	▼ 18.92	▼ 15.45	▼ 21.52	▼ 20.75	▼ 30.00	▼ 21.74	0.00	▼ 26.10	▼ 49.48	-
Exon_x11_7	112:G/A	49.23	0.00	49.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
Exon_x11_7	117:C/T	56.69	0.00	49.23	56.69	0.00	35.29	0.00	41.38	0.00	35.00	-
Exon_x11...	168:T/C	61.11	0.00	33.96	22.50	0.00	61.11	0.00	38.46	0.00	50.00	-
Exon_x11_9	198:C/T	55.74	0.00	0.00	55.74	0.00	0.00	0.00	-	0.00	-	-
Exon_x11...	109:G/T	54.96	0.00	54.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-
Exon_x11...	104:T/C	▼ 19.91	0.00	0.00	0.00	▼ 19.91	0.00	0.00	0.00	0.00	▼ 19.44	-
Exon_x11...	107:C/T	▼ 20.37	0.00	0.00	0.00	▼ 20.37	0.00	0.00	0.00	0.00	▼ 19.44	-
Exon_x11...	265:A/G	54.55	0.00	52.34	51.87	0.00	30.77	0.00	54.55	0.00	27.78	-
Exon_x11...	66:A/G	58.54	0.00	58.54	55.56	0.00	43.33	0.00	45.83	0.00	55.81	-
Exon_x11...	249:C/T	50.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00	-
Exon_x11...	232:A/G	75.00	0.00	45.65	39.33	0.00	75.00	0.00	40.00	0.00	60.00	-
Exon_x11...	125:C/T	▼ 41.79	▼ 27.01	▼ 37.96	▼ 36.75	0.00	0.00	33.33	24.00	▼ 41.79	0.00	-
Exon_x11...	127:T/C	▼ 41.79	▼ 27.01	▼ 38.89	▼ 36.75	0.00	0.00	33.33	24.00	▼ 41.79	0.00	-
Exon_x12	203-204:...	26.42	26.42	16.28	10.71	6.90	0.00	▼ 0.00	0.00	0.00	0.00	-
Exon_x13	196:T/C	64.29	0.00	30.00	47.37	0.00	52.38	0.00	57.14	0.00	64.29	-
Exon_x14	214:A/G	▼ 7.46	0.00	0.00	▼ 7.46	0.00	0.00	0.00	▼ 0.00	0.00	▼ 0.00	-
Exon_x14	215:G/A	▼ 7.46	0.00	0.00	▼ 7.46	0.00	0.00	0.00	▼ 0.00	0.00	▼ 0.00	-
Exon_x16...	231:A/G	59.39	0.00	59.39	44.62	0.00	37.21	0.00	40.43	0.00	49.28	-
Exon_x16...	51:A/G	54.76	0.00	48.85	44.14	0.00	▼ 52.94	0.00	42.86	0.00	54.76	-
Exon_x18	168:A/T	18.18	0.00	0.00	0.00	0.00	0.00	0.00	▼ 0.00	0.00	18.18	-
Exon_x18	169:T/A	18.18	0.00	0.00	0.00	0.00	0.00	0.00	▼ 0.00	0.00	18.18	-
Exon_x18	239:G/A	51.61	0.00	51.61	47.06	0.00	38.89	0.00	▼ 29.17	0.00	50.00	-
Exon_x24	134:C/T	51.43	0.00	0.00	0.00	0.00	0.00	0.00	51.43	0.00	0.00	-
Exon_x10	65:A/T	3.55	0.00	3.55	▼ 3.20	0.00	0.00	0.00	0.00	0.00	0.00	-
Exon_x10	66:T/A	3.55	0.00	3.55	▼ 3.20	0.00	0.00	0.00	0.00	0.00	0.00	-
Exon_x10	80:A/G	1.10	0.00	0.00	0.00	1.10	0.00	0.00	0.00	0.00	0.00	-
Exon_x11_5	38:A/G	2.01	0.00	0.00	1.69	0.00	0.00	0.00	0.00	2.01	1.31	-
Exon_x11_5	125:A/G	1.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.05	-

# Data output - Variant plots



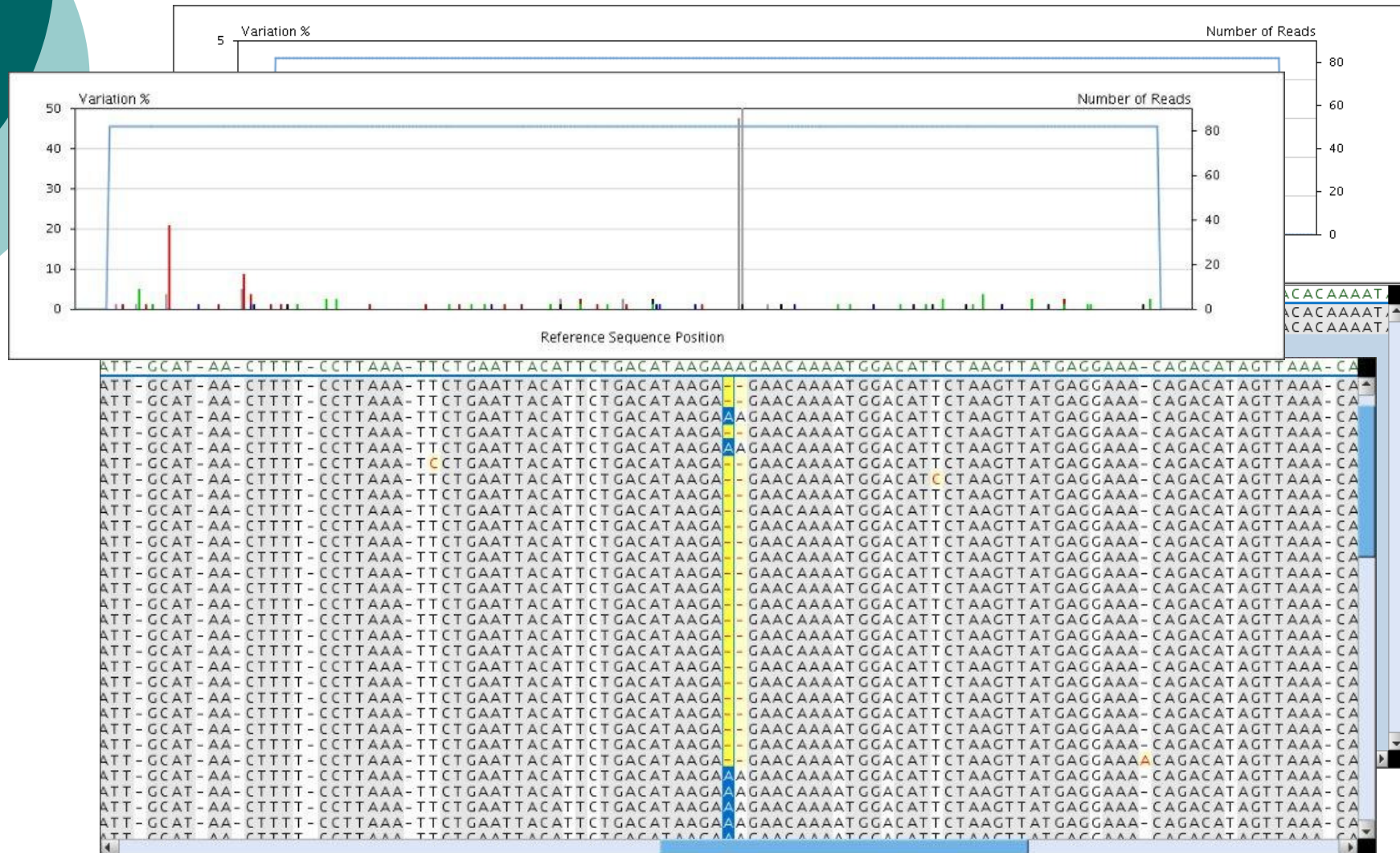
*BRCA2* exon 11 (14)  
 c.4965C>G  
 p.Tyr1655X



*BRCA1* exon 11 (14)  
 c.3770\_3771del  
 p.Glu1257GlyfsX9

# BRCA2 exon 11 (11)4414\_4415delIAA

Not detected in consensus view







## Limits to the amplicon sequencing approach

---

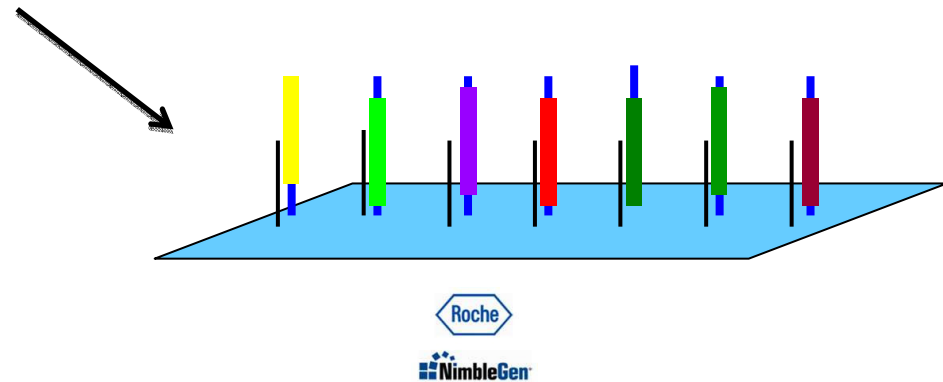
- Management of PCR multiplex
- Alternatives
  - Long range PCR to reduce number of amplicons
  - Sequence capture

# LGMD sequencing

- 150 di/monoplex groups (260 amplicons – 16 genes)

Grp	Gene	Exon	Size	Grp	Gene	Exon	Size	Grp	Gene	Exon	Size	Grp	Gene	Exon	Size
1	DYSF	9	246	27	LMNA	9	348	53	SGCG	3	374	79	CAPN	4	403
1	DYSF	10	246	27	FLNC	26	343	53	VCP	12	368	79	DYSF	33	413
2	ZASP	5	256	28	VCP	4	344	54	FLNC	35	374	80	FLNC	9	418
2	DES	9	261	28	DYSF	7	349	54	ZASP	4a	357	80	FLNC	10	411
3	DES	7	258	29	DYSF	29	344	55	DYSF	14	374	81	FLNC	38	412
3	DYSF	8	269	29	FLNC	34	350	55	FLNC	46a	387	81	DYSF	39	418
4	DYSF	15	260	30	VCP	6	350	56	CAPN	3	374	82	FLNC	21a	424
4	DYSF	32	277	30	DYSF	21	345	56	FLNC	44	387	82	CAV	2b	411
5	DYSF	2	278	31	DYSF	18	346	57	DES	5	387	83	DES	4	427
5	DYSF	31	285	31	FLNC	19	351	57	DYSF	46	375	83	LMNA	6	423
6	CAV	1	280	32	FKBP	1	346	58	CAPN	22	388	84	SGCA	5	421
7	DYSF	40	281	32	DYSF	37	351	58	FLNC	43	393	84	CAPN	11	427
7	DYSF	41	287	33	MYOT	10	359	59	SGCA	1	394	85	MYOT	3	428
8	DYSF	17	281	33	VCP	16	353	59	SGCA	9	388	85	FLNC	17	423
8	DYSF	48	292	34	DYSF	55	359	60	CAPN	7	388	85	MYOT	4	422
9	DYSF	11	296	35	SGCA	6	354	61	SGCG	6	395	86	FLNC	8a	428
9	DYSF	22	291	35	VCP	7	361	61	DYSF	50	389	87	FLNC	39	423
10	DYSF	4	303	35	FLNC	4	361	62	DYSF	19	389	88	FKBP	7	424
10	CAPN	24	308	35	CAPN	17	354	62	DYSF	23	396	88	SGCD	8	430
11	DYSF	44	305	37	DYSF	6	361	63	MYOT	8	395	89	LMNA	6	427
11	DYSF	51	311	37	VCP	14b	354	64	FLNC	3	390	89	FLNC	12	426
12	SGCA	8	313	38	FLNC	7	355	64	LMNA	1b,2	395	90	VCP	2	421
12	CAPN	21	305	38	DYSF	58	354	65	SGCB	5	397	90	SGCB	3	441
13	CAPN	15	315	38	VCP	16	355	65	CAPN	10	392	91	CRYAB	3	428
13	CAPN	18	307	39	DYSF	47	367	66	LMNA	2	397	91	MYOT	7	443
14	FLNC	5	308	40	SGCA	4	357	66	FLNC	11	392	92	CRYAB	1	426
14	FLNC	6	315	40	FLNC	13	368	67	LMNA	3,2	398	92	SGCD	2	441
15	MYOT	5	308	41	SGCG	4	370	67	FLNC	8	392	93	SGCB	6	446
15	FLNC	29	318	41	ZASP	9	357	68	CAPN	9	392	93	CAPN	13	441
16	DES	3	319	42	VCP	5	370	68	DYSF	20	398	94	SGCA	3	447
16	DYSF	16	309	42	DYSF	20	357	69	SGCA	7	399	94	MYOT	2b	442
17	DYSF	28	305	43	FLNC	25	371	69	DYSF	26	404	95	FLNC	28	447
17	LMNA	11b	322	43	FLNC	42	377	70	DES	2	405	95	FLNC	21b	448
18	CAPN	12	323	44	FLNC	24	377	70	SGCD	5	399	96	CAPN	6	442
18	DYSF	53	331	44	FLNC	27	371	71	SGCA	10	400	96	LMNA	7	448
19	SGCG	7	323	45	VCP	8	378	71	FLNC	45	405	97	FLNC	6	449
19	CAPN	8	333	45	ZASP	16	371	72	FLNC	18	405	97	FLNC	41	444
20	DES	8	333	46	LMNA	1b,3	372	72	VCP	13	400	98	SGCD	9	444
21	SGCA	2	334	47	VCP	11	372	73	DYSF	24	407	99	SGCG	8	451
21	SGCD	4	334	47	FLNC	33	380	74	CAPN	19	408	99	DYSF	45	444
22	MYOT	6	339	48	ZASP	15	372	74	FLNC	46b	400	100	LMNA	12,3	445
23	ZASP	5	341	48	CAPN	16	362	75	FKBP	2	400	100	ZASP	4b	453
23	FLNC	25a	385	49	FLNC	32	372	75	ZASP	14	408	101	DES	6	445
24	DYSF	25	336	49	DYSF	32	383	76	SGCG	5	401	101	VCP	14a	453
				50	CAPN	2	383	76	LMNA	11a	409	102	SGCD	6	453

- Sequence capture – 0.8Mb – 26 genes



# Acknowledgements

---



- Jonathan Coxhead
- Charlotte Masters
- Andy Frey



The Development Agency  
for the North East of England



## IT/Bioinformatics

- Katie Siddle
- Dan Swann

The Newcastle upon Tyne Hospitals   
NHS Foundation Trust

## Northern Genetics Service

- Louise Stanley
- James Tellez
- Ciaron McAnulty