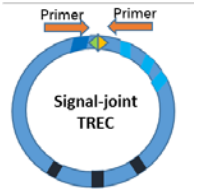
NEW YORK STATE OF OPPORTUNITY. | Department of Health | Wadsworth Center

Design and Validation of a 2nd Tier Next Generation Sequencing (NGS) Panel for Newborn Screening for Severe Combined Immunodeficiency Disease (SCID)

September 13, 2017

Colleen Stevens, Ph.D.
Research Scientist
NYSDOH Wadsworth Center
Newborn Screening Program

Current Testing Algorithm for SCID



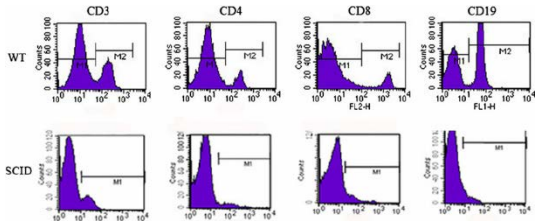
T-cell receptor excision circle (TREC) assay

< 125 Avg. TRECs

Refer to Specialist for Diagnostic Testing

Physician ordered tests

CBC, Flow cytometry, Mitogen studies



Molecular tests – candidate genes; gene panel

- Insurance issues
- Slow TAT (2-8 wks)
- Increased stress to families

NEW YORK STATE Department of Health | Wadsworth Center

Potential Benefits of Molecular Testing by the NBS Program:

- Shortened TAT to identify genetic basis of disease
- Faster diagnosis; phenotype prediction
- Earlier decision on best treatment options – better outcomes
- Cost savings to family and health care system
- Less stress for families

APHL/CDC Cooperative Agreement - Specific Aims

- Develop and validate a 2nd tier multi-gene immunodeficiency panel on 2 different commonly used NGS platforms
- Compare the 2 platforms for accuracy, ease of use, cost and turn-around time
- Evaluate Utility of Targeted NGS for SCID
  - Identify causative gene?
  - Shortened time to diagnosis?
  - Earlier, targeted treatment?
- Provide CDC with quality control/reference samples

NYS NBS 39-gene SCID panel

Gene Selection:

- Commercial SCID panels
- CLSI guideline
- Literature search and case studies

ADA	AK2	ATM	BLNK	BTK	CD3D	CD3E
CD3G	CD247	CD40LG	PTPRC	CHD7	CORO1A	DCLRE1C
DKC1	DOCK2	DOCK8	FOXN1	GATA2	IGHM	IL2RG
IL7R	JAK3	LIG4	MTHFD1	MTR	NHEJ1	NBN
PNP	PRKDC	RAC2	RAG1	RAG2	RMRP	SLC46A1
STAT5B	TBX1	WAS	ZAP70			

NGS Platforms

Illumina MiSeq –
TruSeq Custom
Amplicon (TSCA) panel

Life Technologies
Ion Torrent S5
AmpliSeq (AS) panel

Platform Comparisons

Illumina MiSeq	Specification	Ion Torrent S5
10 ng	Recommended Minimum DNA Input	1 ng
Paired-end reads	Sequence Reads	Single end reads
Ligation capture/amplification	Library Prep/ Gene Targeting	Multiplex PCR
Reversible terminator with fluorescent dNTPs	Sequence detection	Semiconductor based sequencing

MiSeq Flow Cell

The workflow is very similar for both platforms

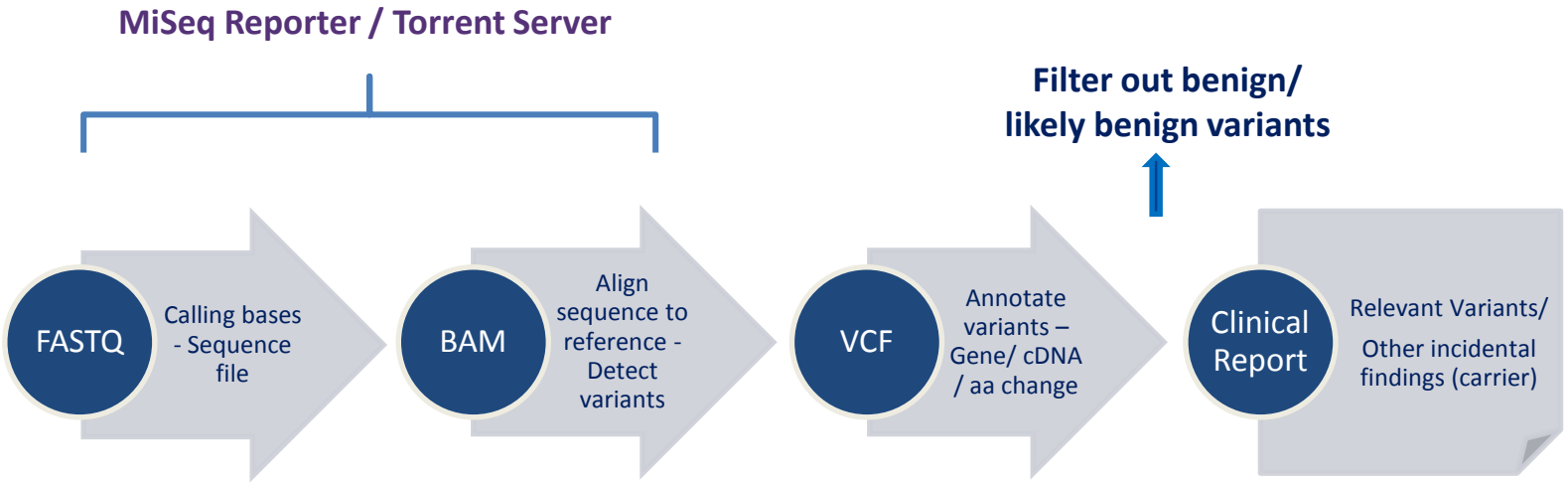
Ion Chip

SCID NGS Turn-Around Time (TAT)

MiSeq		Step	Torrent	
8 Samples	16 Samples		8 Samples	16 Samples
1 hr	1 hr	DNA Extraction (Generations Soln)	1 hr	1 hr
8 hr (3 hrs hands on)	8 hr (3 hrs hands on)	Library Prep *	8 hrs (15 min hands on)	16 hrs (15 min hands on)
15 min	15 min	Loading chip *	15 hrs (15 min hands on)	15 hrs (15 min hands on)
17 hrs	24 hrs	Instrument run time	2.5 hrs	2.5 hrs
4 hrs	6 hrs	Basecalling + alignment (platform software)	5 hrs	8 hrs
30.25 hrs	39.25 hrs	Total TAT (to raw data)	31.5 hrs	42.5 hrs

* Library prep and chip loading: Manual for MiSeq (performed by Core facility)
Ion Chef for Torrent (8 samples at a time)

NGS Data Analysis – Bioinformatics Pipeline

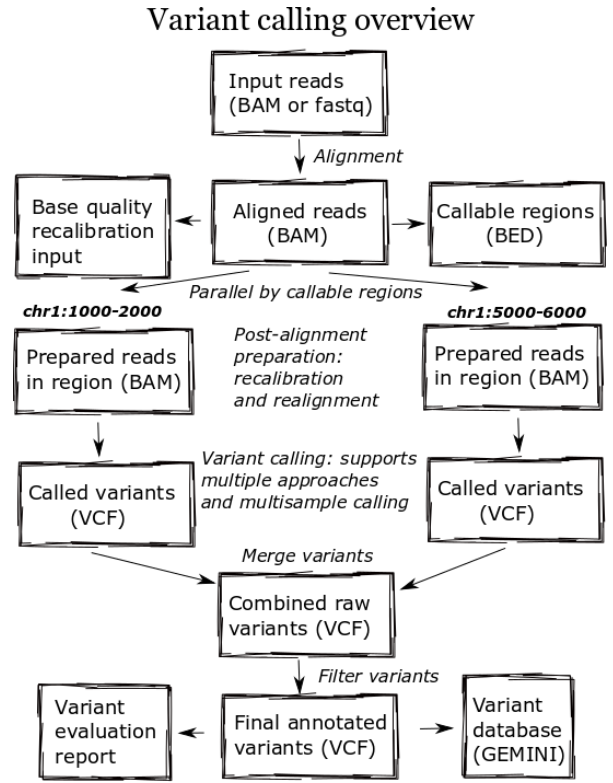


Do you need someone trained in Bioinformatics?

Not necessarily, but…….

Bcbio NGS Bioinformatics Tools

- Genome Build: GRCh37
- Alignment: bwa-mem
- Variant Callers: freebayes, gatk-haplotype & samtools
- Variant Filtering (quality): bcftools
  - RTG vcfeval used to generate thresholds for filtering
- Variant Annotation: snpeff, VEP & vcfanno
  - dbnsfp 3.3a (gnomad, exac, 1000g etc)
  - clinvar 20160502
  - dbsnp 147-20160408
- Variant Prioritization: gemini



https://bcbio-nextgen.readthedocs.io/en/latest/contents/internals.html

Validation Plan

- NGS Sequence
  1) 8 samples from infants diagnosed with SCID – genetic cause known
  2) 16 samples from infants diagnosed with SCID or other immunodeficiency – genetic cause known/unknown
  3) Reference Sample – NA12878 (NIST; GIAB) – variants published
- Sanger sequence
  - 8 samples from #1 above
    - All 39 genes
      - 657 amplicons/sample = 5256 amplicons total
  - 16 samples from #2 above
    - Sanger confirm pathogenic/likely pathogenic and VOUS
- Every sample sequenced in 2 separate runs on each platform – inter-assay reproducibility
- One sample run twice on each run – intra-assay reproducibility

NGS Results for 24 DBS Samples

Metric	TruSeq (MiSeq)	AmpliSeq (Torrent)
Avg. Depth of Coverage	661 (211-884)	586 (247-789)
# Variants called	287 (231-388)	336 (272-384)
% Bases Called (met QC criteria)	98.4% (93-98.8%)	97.8% (95.1 - 98.3%)

Both platforms provide good quality sequence data and adequate coverage of the targeted region

Platform Performance – Example Run

	TruSeq (MiSeq)				AmpliSeq (Ion Torrent)			
Sample	ng DNA	Avg Depth	# Variants	% Uncalled Bases	ng DNA	Avg Depth	# Variants	% Uncalled Bases
101	10	747	351	1.5	5	665	377	1.8
102	10	530	292	1.6	5	789	312	1.8
103	2	211	330	7.0	1.1	487	317	1.9
104	10	461	276	1.4	5	789	303	1.8
105	4	650	337	2.5	2.6	664	385	2.0
106	10	882	269	1.2	5	701	319	1.8
107	10	509	271	1.2	5	776	356	1.7
107_2	10	764	282	1.2	5	770	345	1.7
NA12878	10	622	252	1.4	5	708	307	1.7
NA12878_low	2	529	250	3.2	1.1	489	328	1.8

AmpliSeq appears somewhat more tolerant to low DNA concentration

Accuracy – Reference Sample NA12878

Panel	Run	Sample	FN	FP	TP	TN	Sensitivity	Specificity
TruSeq (MiSeq)	1	NA12878	0	3	76	135581	1	0.9999
	2	NA12878	2	5	74	134847	0.9737	0.9999
AmpliSeq (Ion Torrent)	1	NA12878	2	0	74	134732	0.9737	0.9999
	2	NA12878	3	2	73	135581	0.9605	0.9999

- Both platforms demonstrated excellent sensitivity and specificity for identification of variants in the reference sample
- Results were reproducible between runs

8 Validation Samples – Confirmed Genetic Cause

Sample	Gene	Variant (cDNA)	Variant (aa)	Zygosity	Pathogenicity Classification	Verified
102	CHD7	c.434G>A	p.Trp145Ter	Het (dom)	Pathogenic	Yes
105	ADA	c.301C>T	p.Arg101Trp	Hom	Pathogenic	Yes
108	IL7R	c.83-2A>T	-	Het	VOUS	Yes
	IL7R	c.353G>A	p.Cys118Tyr	Het	Pathogenic	
110	IL2RG	c.982C>T	p.Arg328Ter	Hom	Likely Pathogenic	Yes
113	IL2RG	c.865C>T	p.Arg289Ter	Hom	Pathogenic	Yes
115	ADA	c.301C>T	p.Arg101Trp	Hom	Pathogenic	Yes
116	JAK3	c.2872G>T	p.Glu958Ter	Het	Likely Pathogenic	Yes
		c.1261delC	p.Leu421fs	Het	Likely Pathogenic	
120	RAG2	c.501A>C	p.Arg167Ser	Het	VOUS	Yes
		c.283G>A	p.Gly95Arg	Het	Pathogenic	

Variants filtered for pathogenic/likely pathogenic - homozygous or compound heterozygote. VOUS unfiltered if 1 pathogenic/likely pathogenic.

16 Validation Samples – Genetic Cause Known/Unknown

* Genetic cause known

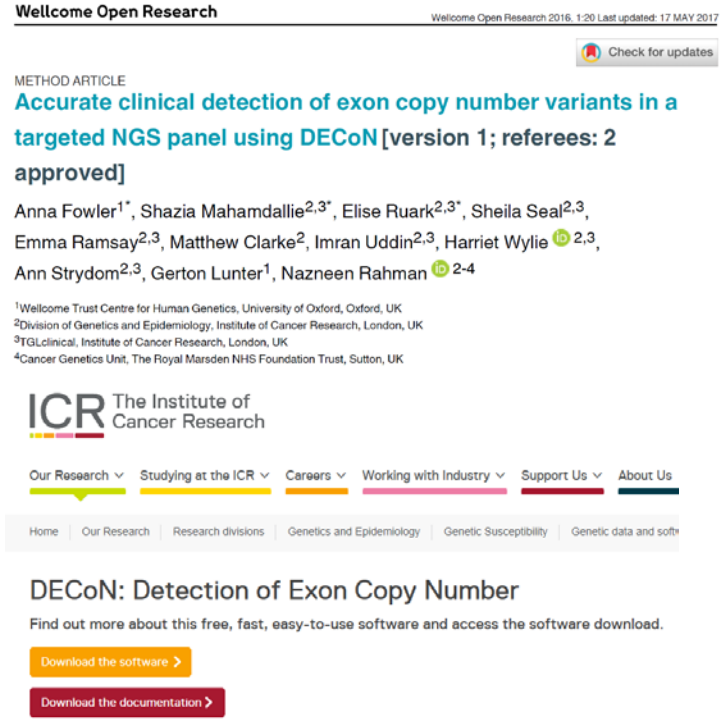
Sample	Gene	Variants	Sample	Gene	Variants
101	RAG1	p.Trp522Cys;p.Arg778Trp	114	None	
103*	ADA	p.Pro297del (Hom)	117	None	
104	None		118	ATM	c.5763-1G>A (Hom)
106	ADA	p.Arg101Trp (Hom)	119	IL2RG	p.Asp194Tyr (Hom)
107	None		121*	None	
109*	None		122	None	
111	ADA	p.Gly216Arg (Hom)	123	RAG1	p.Arg507Trp;p.Ser966Thr
112*	None		124*	IL2RG	p.Leu272fs (Hom)

Further analysis to be done:

- Sanger confirm all pathogenic/likely pathogenic/VOUS
- Re-assess potentially causative variants – 2 VOUS

DECoN

Uses depth of coverage to copy number variants



Wellcome Open Research

Wellcome Open Research 2016, 1:20 Last updated: 17 MAY 2017

Check for updates

METHOD ARTICLE

Accurate clinical detection of exon copy number variants in a targeted NGS panel using DECoN [version 1; referees: 2 approved]

Anna Fowler[1*], Shazia Mahamdallie[2,3*], Elise Ruark[2,3*], Sheila Seal[2,3], Emma Ramsay[2,3], Matthew Clarke[2], Imran Uddin[2,3], Harriet Wylie[2,3], Ann Strydom[2,3], Gerton Lunter[1], Nazneen Rahman[2-4]

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[2]Division of Genetics and Epidemiology, Institute of Cancer Research, London, UK
[3]TGLclinical, Institute of Cancer Research, London, UK
[4]Cancer Genetics Unit, The Royal Marsden NHS Foundation Trust, Sutton, UK

ICR The Institute of Cancer Research

Our Research | Studying at the ICR | Careers | Working with Industry | Support Us | About Us

Home | Our Research | Research divisions | Genetics and Epidemiology | Genetic Susceptibility | Genetic data and soft

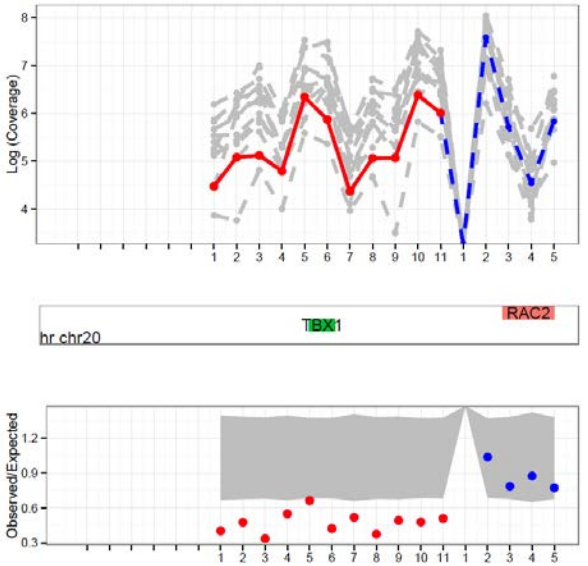
DECoN: Detection of Exon Copy Number

Find out more about this free, fast, easy-to-use software and access the software download.

Download the software

Download the documentation

Sample 121



Samples 109 and 112 also had low coverage for TBX

These 3 infants were reported as DiGeorge Syndrome (heterozygous deletion of TBX1)

NEW YORK STATE | Department of Health | Wadsworth Center

Summary

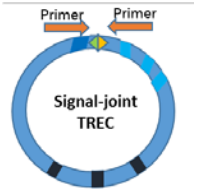
- TruSeq (MiSeq) and AmpliSeq (Ion Torrent) 39-gene SCID panels demonstrate similar accuracy and reproducibility for detection of variants in the reference sample NA12878
- Both panels detected known pathogenic variants in infants diagnosed with SCID
- Potential genetic causes were identified for infants diagnosed with an immunodeficiency for whom we do not have genetic information.

Next Steps

- Complete validation
  - Sanger confirmations
  - Sensitivity and Specificity of assay for dried blood spot samples
  - Submit validation to NYS Clinical Laboratory Evaluation Program for approval

- Begin Consented Study

Proposed Testing Algorithm for SCID

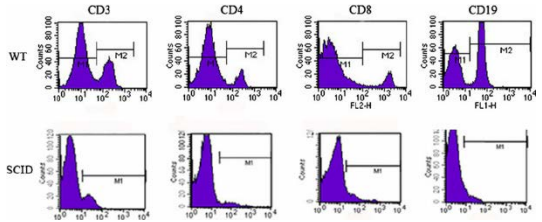


T-cell receptor excision circle (TREC) assay

< 125 Avg. TRECs

Refer to Specialist for Diagnostic Testing

CBC, Flow cytometry, Mitogen studies



Immunodeficiency identified: Informed Consent

NBS NGS: 39-gene SCID panel

- No Insurance issues
- TAT (~1 wk)
- Less stress on families

NEW YORK STATE Department of Health | Wadsworth Center

Acknowledgments:

NYS NBS Program:
Michele Caggana, ScD, FACMG
Bob Sicko, BS
Allison Madole, BS
Denise Kay, PhD
Carlos Saavedra, MD
Beth Vogel, MS, CGC
Sarah Bradley, MS, CGC

Wadsworth Center Applied Genomic Technologies Core