ENDIA

Serology & virology division
diagnosis.research.teaching

SEALS
excellence in diagnosis research teaching

UNSW AUSTRALIA

NGS and Virus Discovery

Viruses in May 2018

Never Stand Still | Medicine | School of Women's & Children's Health

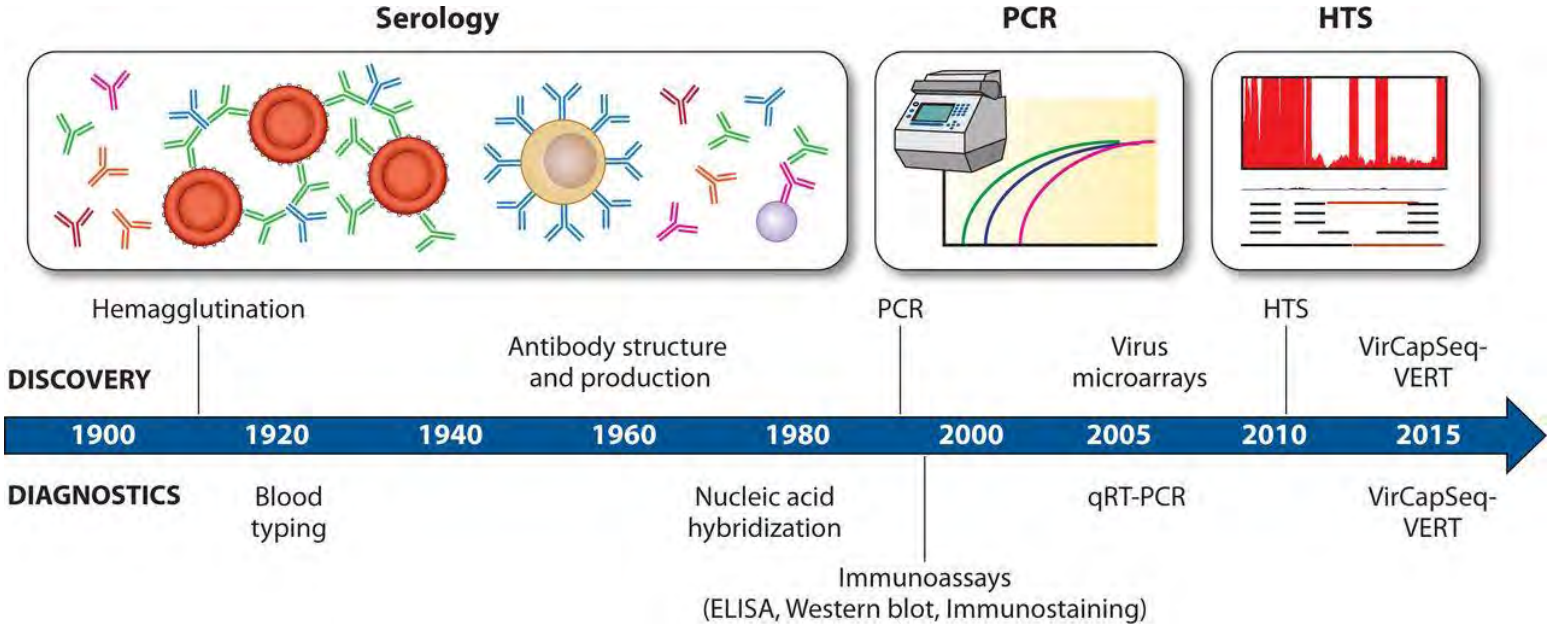
Dr. Ki Wook Kim

UNSW Virology Research Laboratory
Prince of Wales Hospital
School of Women's and Children's Health,
UNSW

Contents

- Evolution of virus detection
- Human virome
- Bottlenecks of virome sequencing
- Virome-capture sequencing
- Application in type 1 diabetes research
- Viruses in the Genetically at Risk (VIGR)
- Environmental Determinants of Islet Autoimmunity (ENDIA)

Evolution of virus detection

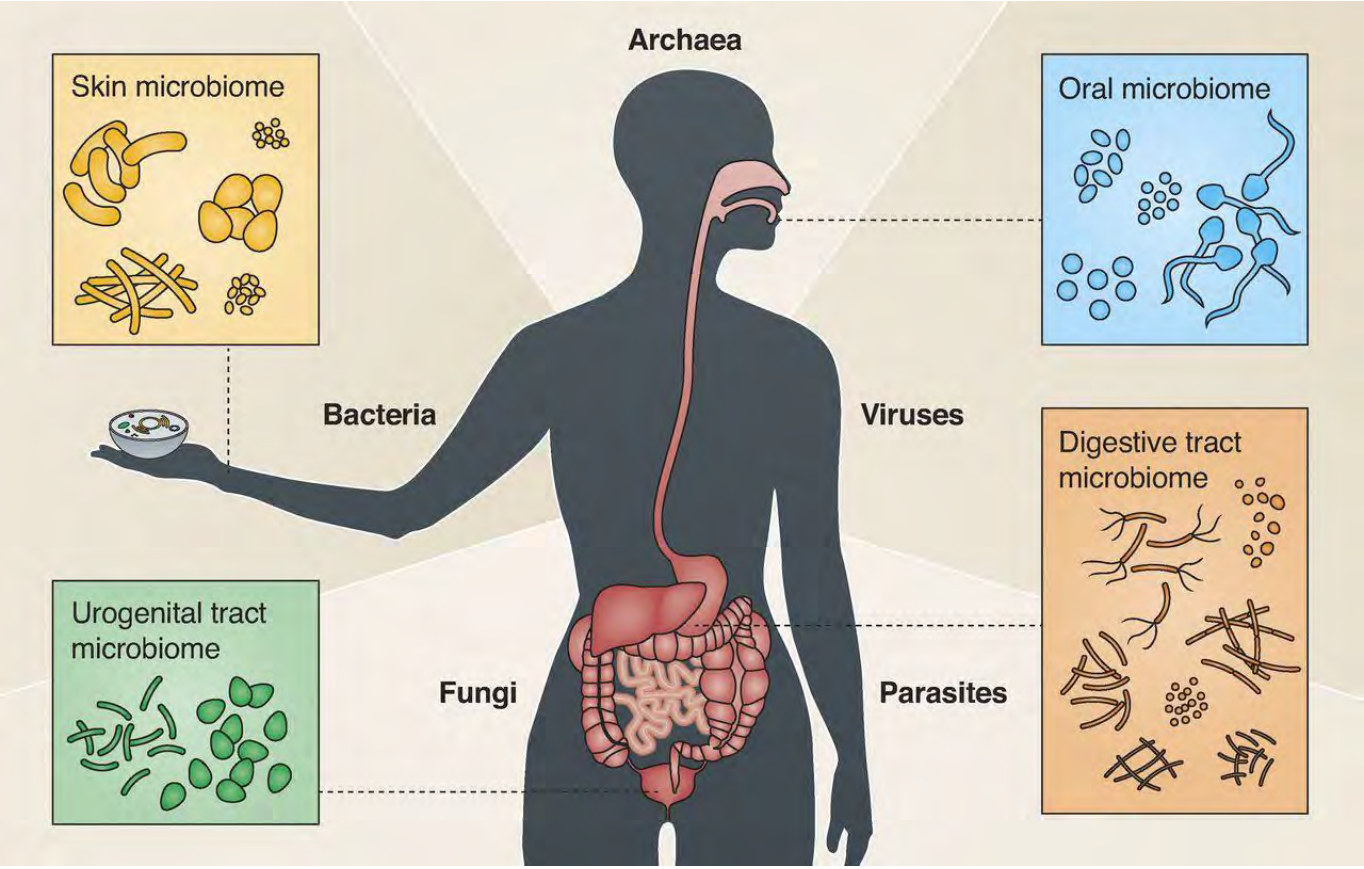


Rasmussen *mBio* 2015

Wouldn't it be nice if we could…

- Simply sequence them all? i.e. the "virome"
- Culture-independently
- No *a priori* knowledge of viruses in the sample
- Aspiration since advent of NGS (>10 yrs)

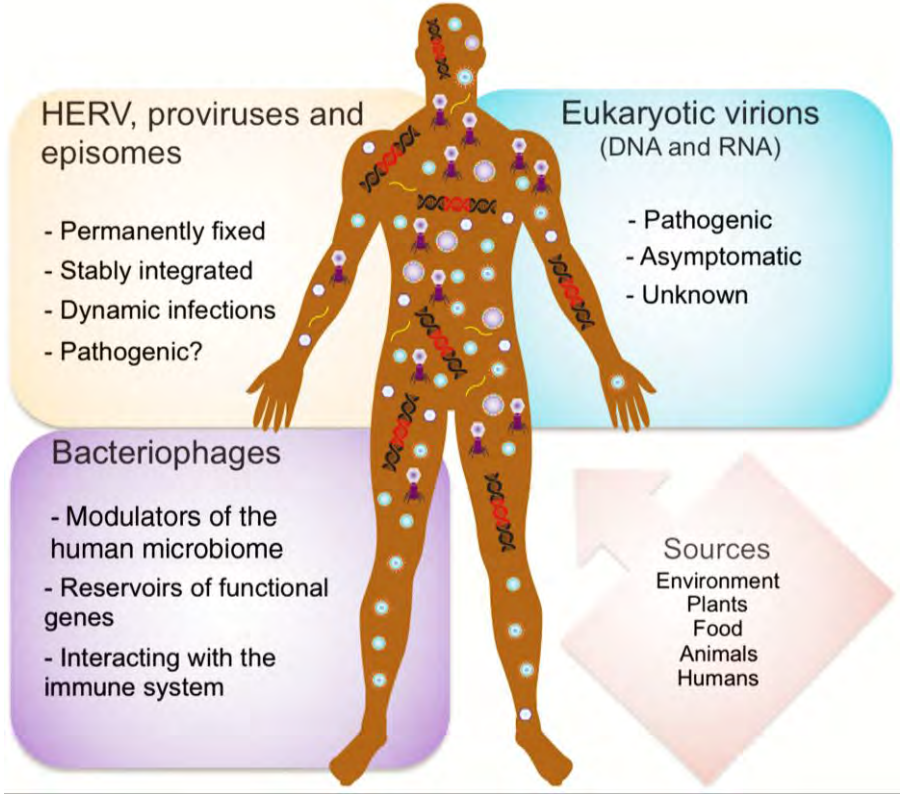
Human microbiome



Garrette, *J Cell Biol* 2015

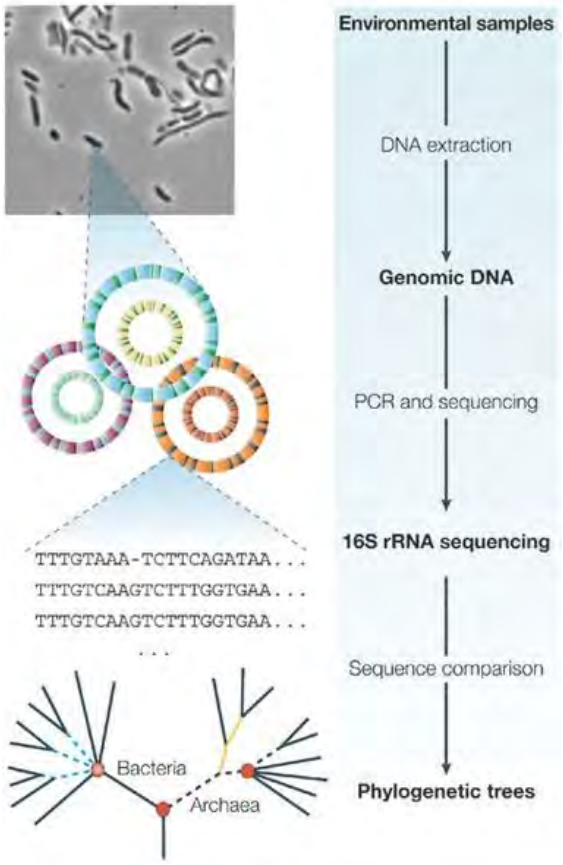
Human virome

- Vertebrate-infecting viruses most relevant to disease pathogenesis



Forterre, *Clin Infect Dis.* 2017

No 16S rRNA in viruses



http://www.nxt-dx.com/metagenomics/16s-rrna-sequencing/

Bottlenecks of virome NGS

- No 16S equivalent
- >99% reads bacteria or host
- Majority of viral sequences from phage
- Most RNA viruses in stool are plant viruses
- **Needle in a haystack!**

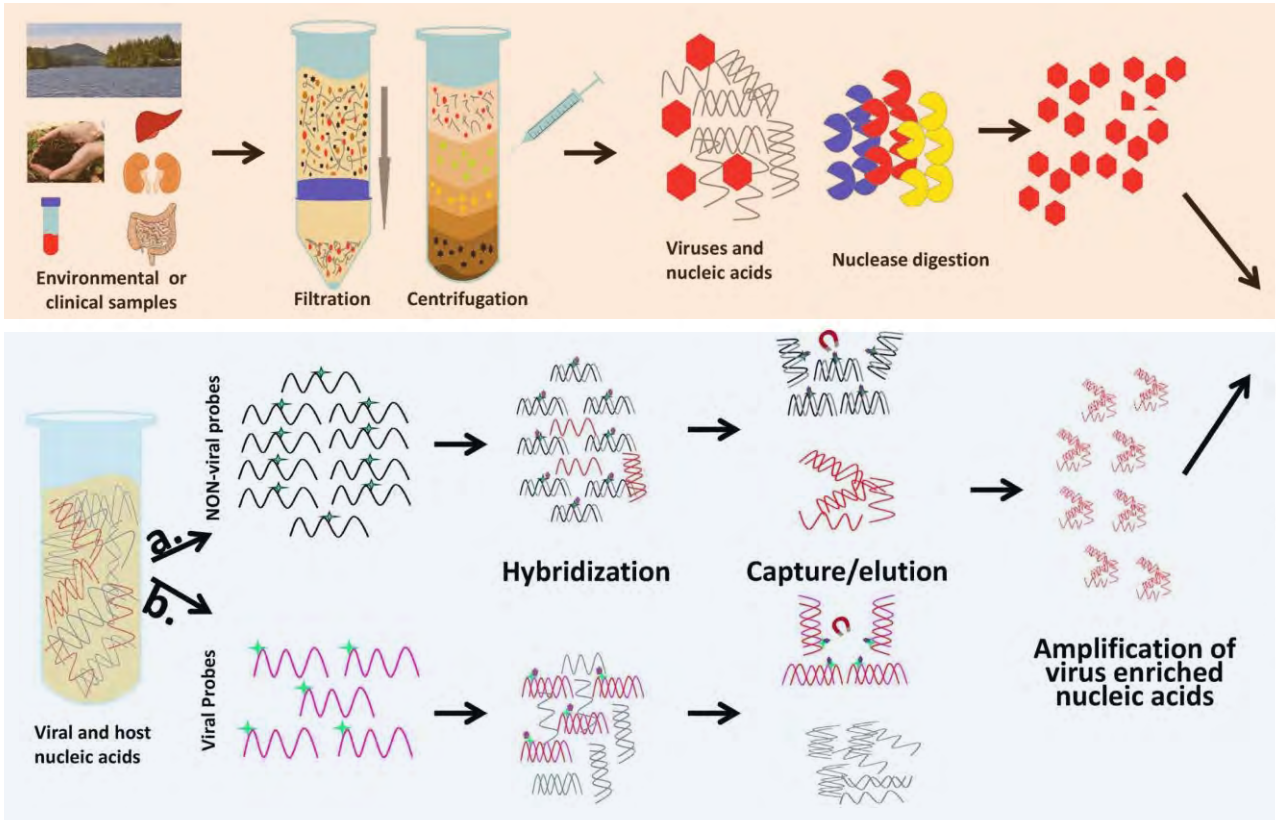
Viral sequence enrichment strategies

Physical enrichment

Sequence capture

a.Depletion

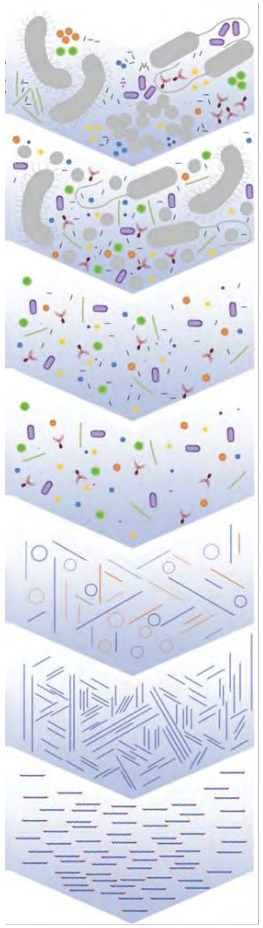
b.Enrichment



Adapted from Kumar et al., *Virus Research* 2017

Conventional virome NGS

- Physical enrichment
- 2- or 3-step process
- Sample tampering
- Loss of viral NA
- Laborious
- Slow
- Expensive
- ~20%↑ in viral reads



Homogenisation

Ultra-centrifugation & Filtration

Nuclease treatment
Assumption that viral NA protected by necleocapsid

Ribo-zero
Human rRNA depletion $$$

DNA/RNA purification

Random PCR amplification

NGS Library Preparation

Illumina Sequencing

Adapted from Conceição-Neto et al., *Sci Reports* 2015

Sequence Capture enrichment for NGS

- Widely applied for whole exome sequencing (Cancer genetics)
  - 30 Mb vs 3,200 Mb (Human genome)
- Biotinylated ssRNA/DNA baits (50-150 mer)
- Solution-based capture using magnetic streptavidin beads

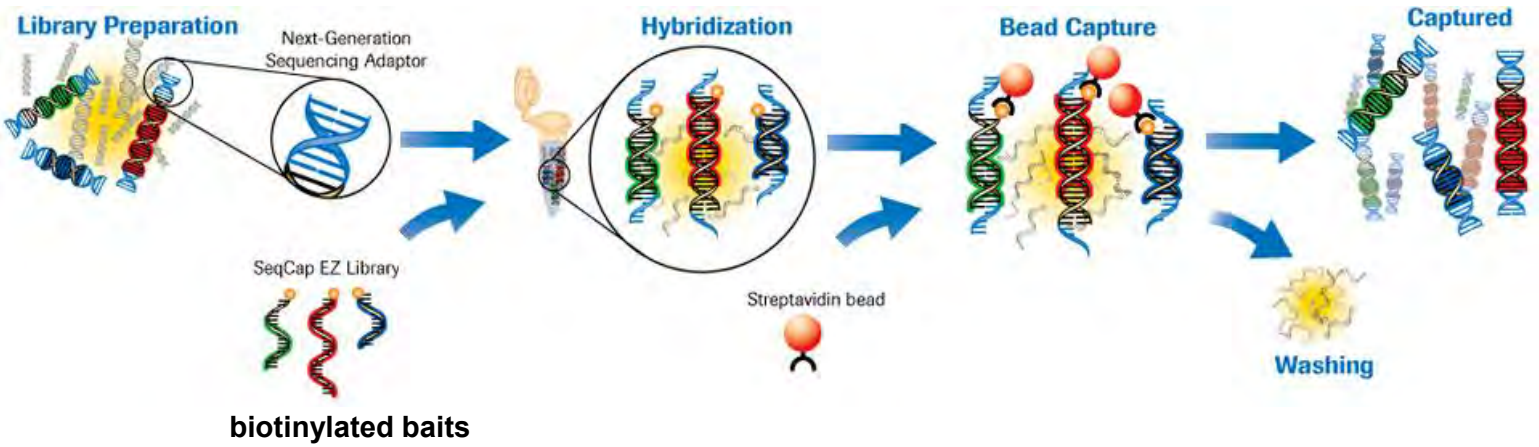


Table 1
Exome capture technology designs

	NimbleGen SeqCap EZ	Agilent SureSelect	Illumina TruSeq	Illumina Nextera
Bait type	DNA	RNA	DNA	DNA
Bait length range (bp)	NP	114-126	95	95
Median bait length (bp)	NP	119	95	95
Number of baits	NP	554,079	347,517	347,517
Total bait length (Mb)	NP	66.48	33.01	33.01
Target length range (bp)	59–742	114–21,747	2–37,917	2–37,917
Median target length (bp)	171	200	135	135
Number of targets	368,146	185,636	201071	201,071
Total target length (Mb)	64.19	51.18	62.08	62.08
Fragmentation method	Ultrasonication	Ultrasonication	Ultrasonication	Transposomes
Automation	++	++	++	+++
Throughput	+++	+++	+++	+++
Flexibility	Custom available	Custom available		Custom available
Species	Human, mouse, 3 plant species	Human, mouse, 14 other species custom	Human	Human
Costs	$$	$$	$	$

- xGen (IDT)
- Nugen Ovation
- NEBNext Direct®
- Qiagen

Chilamakuri et al., *BMC Genomics* 2014

VZV, EBV, KSHV

Specific Capture and Whole-Genome Sequencing of Viruses from Clinical Samples

Daniel P. Depledge, Anne L. Palser, Simon J. Watson, Imogen Yi-Chun Lai, Eleanor R. Gray, Paul Grant, Ravinder K. Kanda, Emily Leproust, Paul Kellam, Judith Breuer

Published: November 18, 2011 • https://doi.org/10.1371/journal.pone.0027805

Retroviruses

Application of targeted enrichment to next-generation sequencing of retroviruses integrated into the host human genome

Paola Miyazato, Hiroo Katsuya, Asami Fukuda, Yoshikazu Uchiyama, Misaki Matsuo, Michiyo Tokunaga, Shinjiro Hino, Mitsuyoshi Nakao & Yorifumi Satou

Scientific Reports **6**, Article number: 28324 | Received: 02 April 2016

HCV, HIV

ve-SEQ: Robust, unbiased enrichment for streamlined detection and whole-genome sequencing of HCV and other highly diverse pathogens [version 1; referees: 2 approved, 1 approved with reservations]

David Bonsall[1*], M. Azim Ansari[1,2*], Camilla Ip[3*], Amy Trebes[3], Anthony Brown[1], Paul Klenerman[1,4], David Buck[3], STOP-HCV Consortium, Paolo Piazza[3], Eleanor Barnes[1,4], Rory Bowden[3]

Norovirus

Norovirus Whole-Genome Sequencing by SureSelect Target Enrichment: a Robust and Sensitive Method

Julianne R. Brown,[a,b] Sunando Roy,[c] Christopher Ruis,[c] Erika Yara Romero,[c] Divya Shah,[a,b] Rachel Williams,[c] Judy Breuer[a,c]

Microbiology, Virology, and Infection Control, Great Ormond Street Hospital for Children NHS Foundation Trust, London, United Kingdom[a]; NIHR Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London, London, United Kingdom[b]; Division of Infection and Immunity, University College London, London, United Kingdom[c]

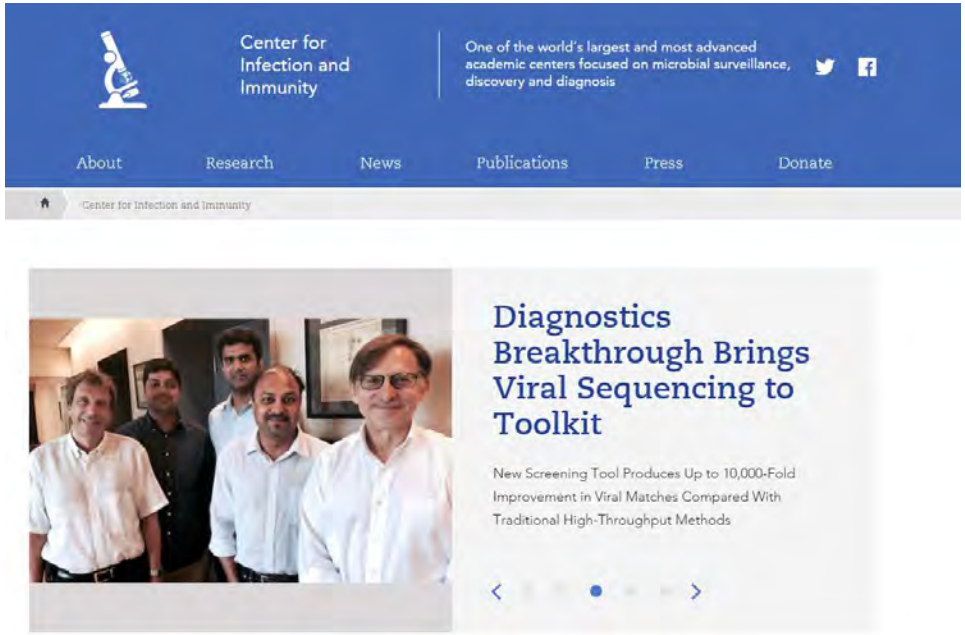
UNSW AUSTRALIA

Table 1
Exome capture technology designs

	NimbleGen SeqCap EZ	Agilent SureSelect	Illumina TruSeq	Illumina Nextera
Bait type	DNA	RNA	DNA	DNA
Bait length range (bp)	NP	114-126	95	95
Median bait length (bp)	NP	119	95	95
Number of baits	NP	554,079	347,517	347,517
Total bait length (Mb)	NP	66.48	33.01	33.01
Target length range (bp)	59–742	114–21,747	2–37,917	2–37,917
Median target length (bp)	171	200	135	135
Number of targets	368,146	185,636	201071	201,071
Total target length (Mb)	64.19 200	51.18	62.08	62.08
Fragmentation method	Ultrasonication	Ultrasonication	Ultrasonication	Transposomes
Automation	++	++	++	+++
Throughput	+++	+++	+++	+++
Flexibility	Custom available	Custom available		Custom available
Species	Human, mouse, 3 plant species	Human, mouse, 14 other species custom	Human	Human
Costs	$$	$$	$	$

Chilamakuri et al., *BMC Genomics* 2014

Virome Capture Sequencing for vertebrate-infecting viruses (VirCapSeq-VERT)



~2 million probes (avg. 70 mer)

complete genomes of all known vertebrate-infectious viruses (207 viral taxa)

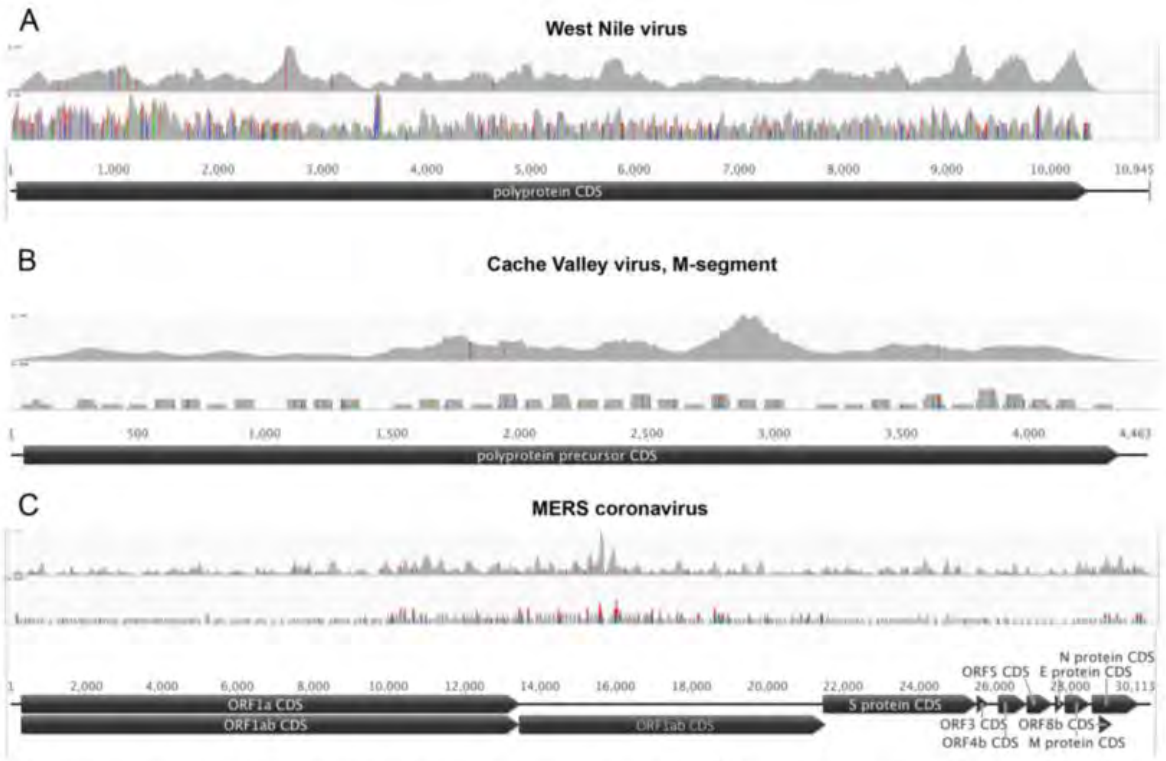
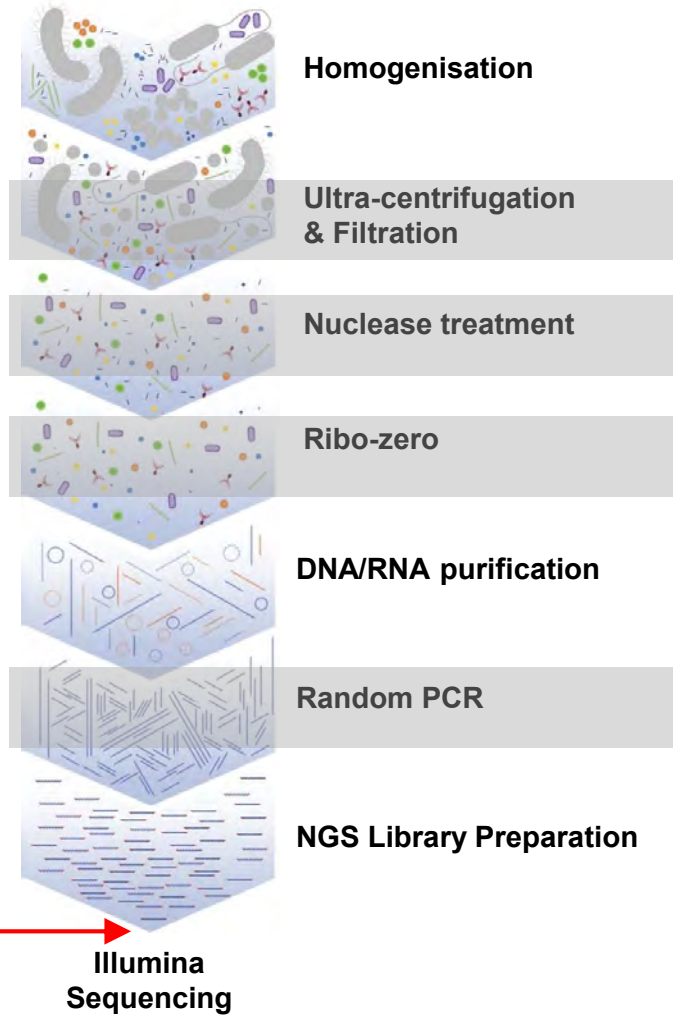


FIG 3 Read coverage versus probe coverage of VirCapSeq-VERT for West Nile virus (A), Cache Valley virus (B), and MERS coronavirus (C). Virus genomes are represented by horizontal black lines and coding sequence by black pointed boxes. The top graph in each panel indicates the read coverage obtained by VirCapSeq-VERT; probe coverage is shown below. Colored lines indicate mismatch to the reference used for alignment (green, A; red, T; blue, C; orange, G). Line heights indicate the frequency of the mismatched bases.

Briese et al., *mBio* 2015

VirCapSeq-VERT

- **Up to 10,000-fold** increase in viral reads
- More sensitive than qPCR
- No special equipment or enzymes
- Post-library prep enrichment
- All known vertebrate-infecting virus



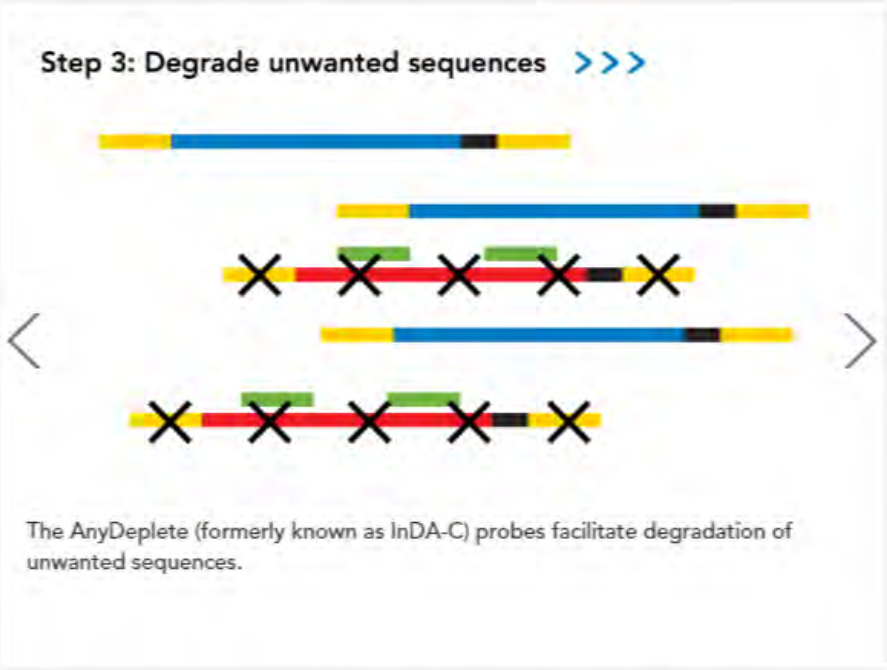
- Can detect novel viruses with up to ~40% sequence divergence

Virome Analysis of Transfusion Recipients Reveals a Novel Human Virus That Shares Genomic Features with Hepaciviruses and Pegiviruses

Amit Kapoor[a], Arvind Kumar[a], Peter Simmonds[b], Nishit Bhuva[a], Lokendra Singh Chauhan[a], Bohyun Lee[a], Amadou Alpha Sall[c], Zhezhen Jin[d], Stephen S. Morse[e], Beth Shaz[f], Peter D. Burbelo[g], W. Ian Lipkin[a]

NGS and Virus Discovery

AnyDeplete (formerly InDA-C)



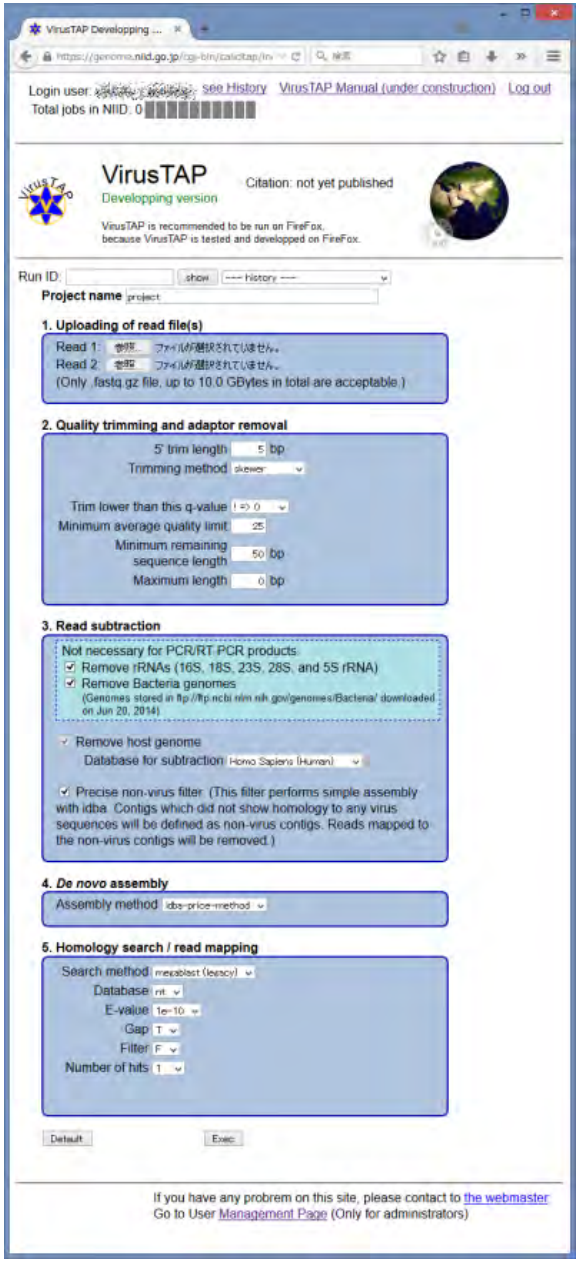
The AnyDeplete (formerly known as InDA-C) probes facilitate degradation of unwanted sequences.

AnyDeplete (formerly InDA-C) is a technology for targeted depletion of abundant transcripts.

- ✓ Customizable to any transcript, any organism
- ✓ Post library depletion avoids off-target mRNA cross-hybridization events that can introduce unwanted bias
- ✓ Maximize informative sequencing reads in whole transcriptome data

NuGEN

UNSW AUSTRALIA


VirusTAP Developping ...
Login user:
see History
VirusTAP Manual (under construction)
Log out
Total jobs in NIID: 0
VirusTAP
Developping version
Citation: not yet published
VirusTAP is recommended to be run on FireFox,
because VirusTAP is tested and developped on FireFox.
Run ID:
show
--- history ---
Project name
project
1. Uploading of read file(s)
Read 1:
Read 2:
(Only .fastq.gz file, up to 10.0 GBytes in total are acceptable.)
2. Quality trimming and adaptor removal
5' trim length
bp
Trimming method
skewer
Trim lower than this q-value
Minimum average quality limit
Minimum remaining sequence length
Maximum length
3. Read subtraction
Not necessary for PCR/RT-PCR products
Remove rRNAs (16S, 18S, 23S, 28S, and 5S rRNA)
Remove Bacteria genomes
(Genomes stored in ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/ downloaded on Jun 20, 2014)
Remove host genome
Database for subtraction
Homo Sapiens (Human)
Precise non-virus filter (This filter performs simple assembly with idba. Contigs which did not show homology to any virus sequences will be defined as non-virus contigs. Reads mapped to the non-virus contigs will be removed.)
4. De novo assembly
Assembly method
idba-price-method
5. Homology search / read mapping
Search method
megablast (legacy)
Database
nt
E-value
1e-10
Gap
T
Filter
F
Number of hits
1
Default
Exec
If you have any probrem on this site, please contact to the webmaster
Go to User Management Page (Only for administrators)

Contents

- Evolution of virus detection
- Human virome
- Bottlenecks of virome sequencing
- Virome-capture sequencing
- Application in type 1 diabetes research
- Viruses in the Genetically at Risk (VIGR)
- Environmental Determinants of Islet Autoimmunity (ENDIA)

Q. Why viruses?

Genetic risk factors for type 1 diabetes

- > 50 genetic loci associated with T1D

Nat Genet. 2009 Jun;41(6):703-7. doi: 10.1038/ng.381. Epub 2009 May 10.

Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes.

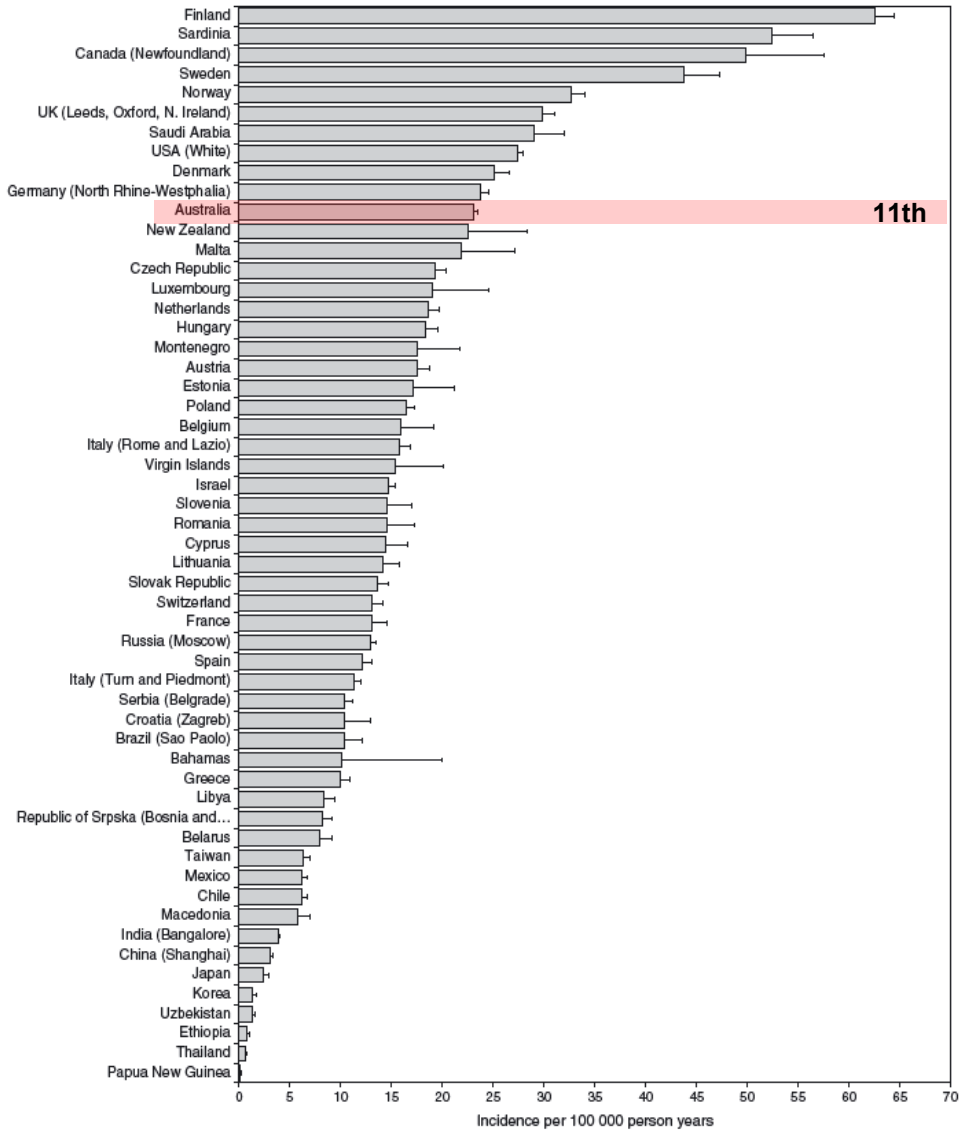
Barrett JC[1], Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS; Type 1 Diabetes Genetics Consortium.

PLoS Genet. 2011 Sep;7(9):e1002293. doi: 10.1371/journal.pgen.1002293. Epub 2011 Sep 29.

A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci.

Bradfield JP[1], Qu HQ, Wang K, Zhang H, Sleiman PM, Kim CE, Mentch FD, Qiu H, Glessner JT, Thomas KA, Frackelton EC, Chiavacci RM, Imielinski M, Monos DS, Pandey R, Bakay M, Grant SF, Polychronakos C, Hakonarson H.

- Human Leukocyte Antigen (HLA) region account ~50% of reported genetic risk



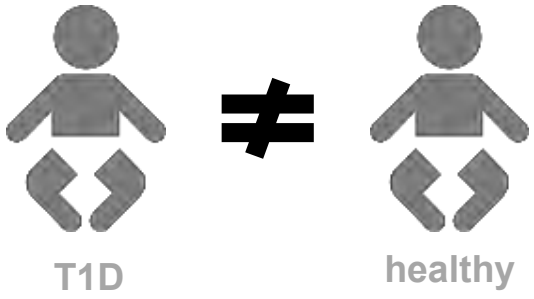
- Since 1980s childhood T1D has **DOUBLED**

- In children (0-14 yrs):
  - **7th** highest prevalence
  - **6th** highest incidence

Craig *Pediatric Diabetes* 2014
Australian Institute of Health and Welfare 2010-2011

Increasing role of the environment

- Significant geographical variation
- Seasonal variation
- Monozygotic twins discordant for T1D



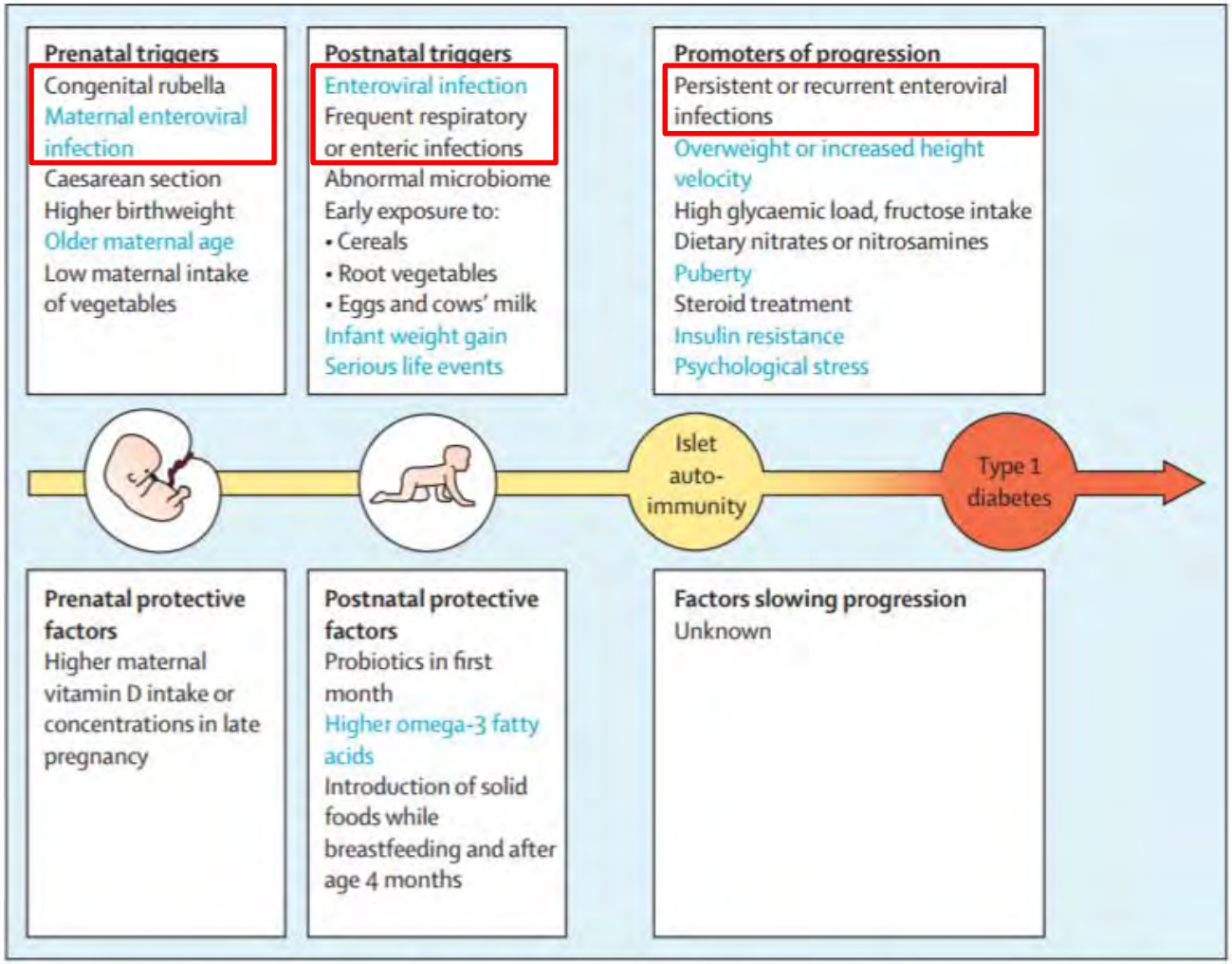


Figure 1: **Environmental triggers and protective factors for islet autoimmunity and promoters of progression to type 1 diabetes for which an association has been suggested**
Triggers and factors with the strongest evidence base are shown in blue.

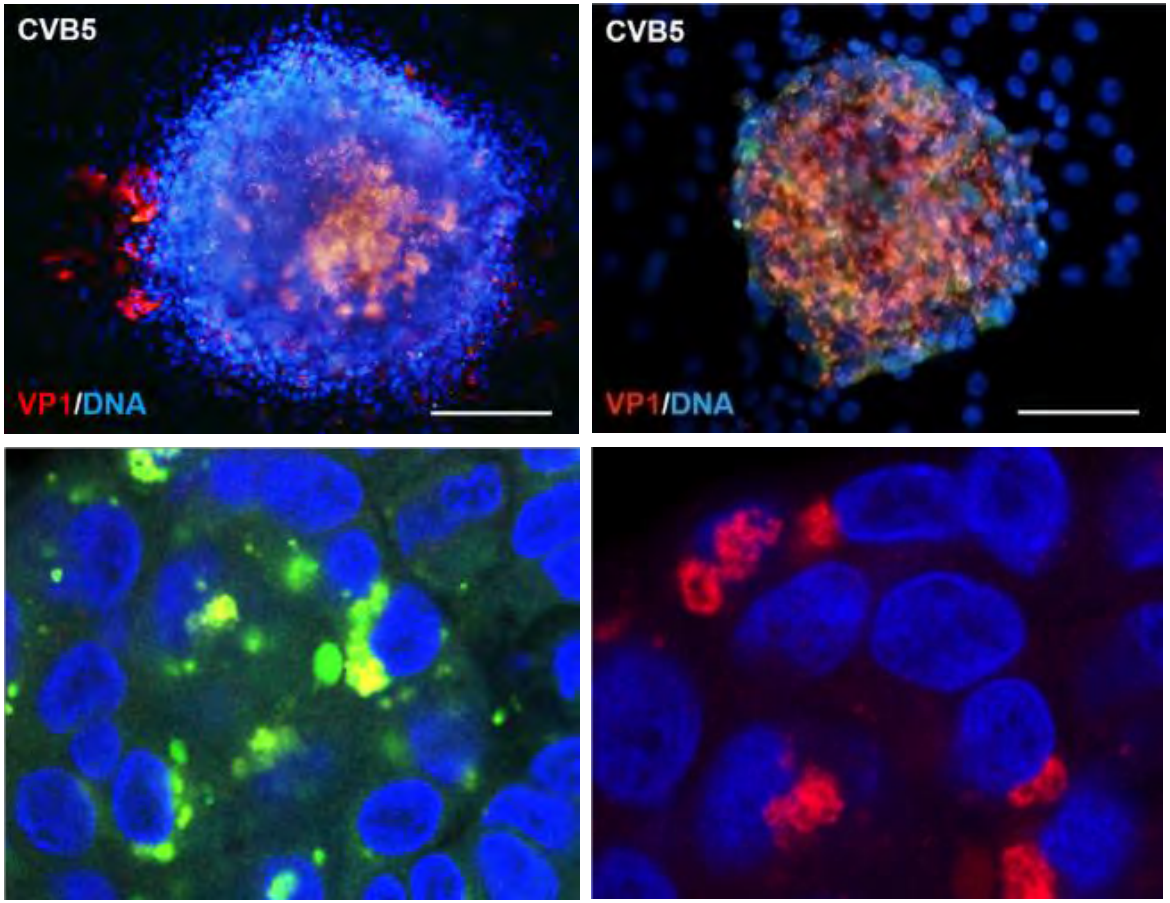
Rewers and Ludvigsson, *Lancet* 2016

UNSW AUSTRALIA

Viral aetiology of T1D

- Prime environmental trigger of islet autoimmunity (IA)
- IA precedes most T1D during childhood
- Strong epidemiological association:
  - EV and IA (odds ratio 3.7)
  - EV and T1D (odds ratio 9.8)
- Detection of EV RNA /protein in pancreata of recent onset patients
- coxsackievirus B implicated and examined most extensively

Yeung et al., *BMJ* 2011
Krogvold et al., *Diabetes* 2014



Insulin

VP1

Why examine the virome?

- Selective focus on EVs over others
  - E.g. rotavirus, rubella, cytomegalovirus and influenza
- Reliance on targeted virus detection (PCR/serology)
- Other viruses may have been missed
- Experimental bias → overestimated association between EVs and T1D?
- Defining the full breadth diabetogenic viruses paramount for vaccine development

Study population

Viruses In the Genetically at Risk (VIGR)

- 93 children with 1st degree relative with T1D (parent or sibling)
- Gut virome: 20 IA+ cases vs IA- (n=64)
- Plasma virome: 41 IA+ cases vs IA- (n=118)

Hypothesis & anticipated outcomes

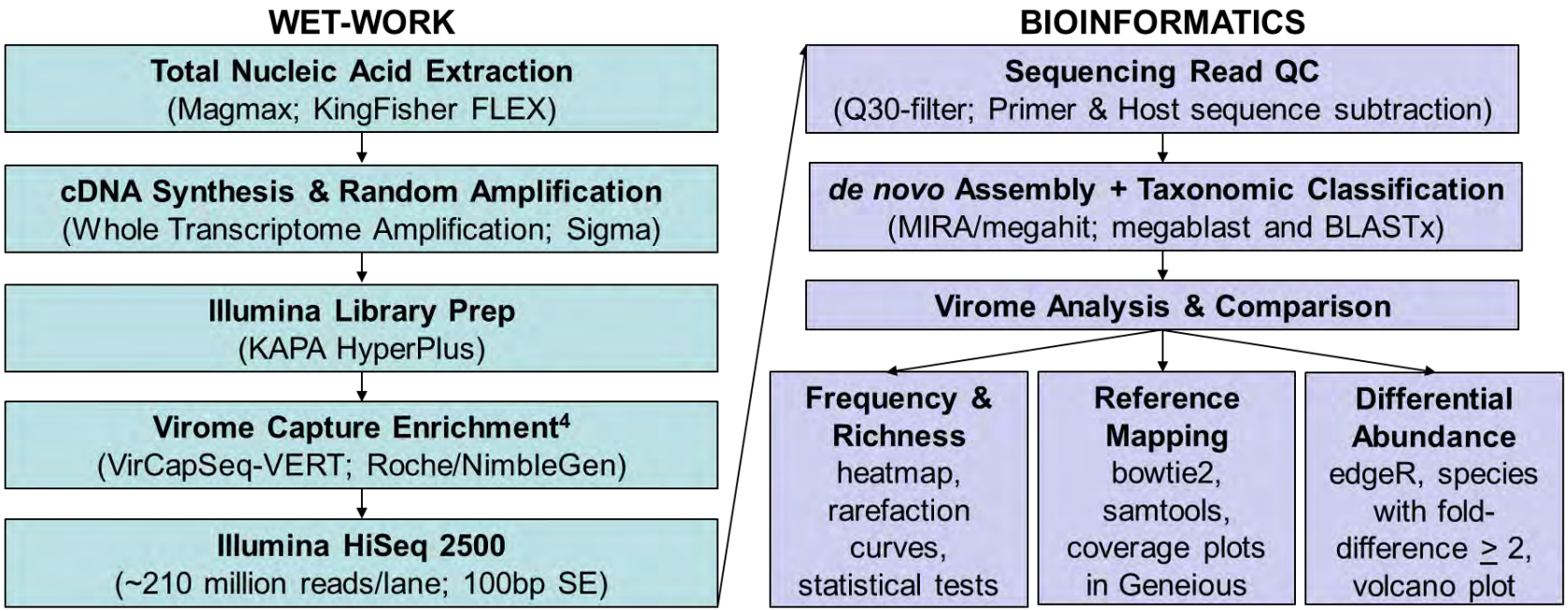
Overarching hypothesis:

- Viral infection in early life increases risk of IA in genetically predisposed children

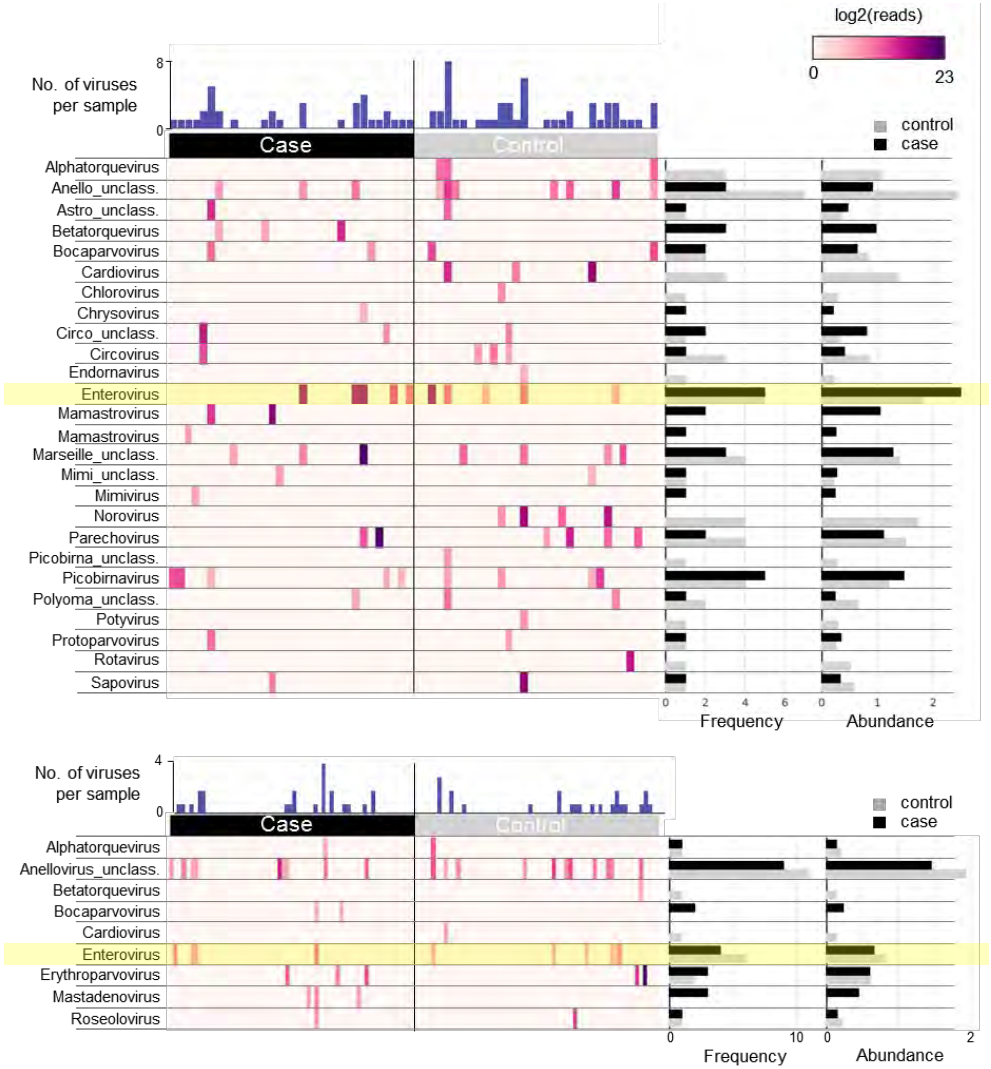
Anticipated outcomes:

- Virome of IA+ cases differ from IA- controls
- EV infections more frequent in IA+ cases

Methods


WET-WORK
Total Nucleic Acid Extraction
(Magmax; KingFisher FLEX)
cDNA Synthesis & Random Amplification
(Whole Transcriptome Amplification; Sigma)
Illumina Library Prep
(KAPA HyperPlus)
Virome Capture Enrichment[4]
(VirCapSeq-VERT; Roche/NimbleGen)
Illumina HiSeq 2500
(~210 million reads/lane; 100bp SE)
BIOINFORMATICS
Sequencing Read QC
(Q30-filter; Primer & Host sequence subtraction)
de novo Assembly + Taxonomic Classification
(MIRA/megahit; megablast and BLASTx)
Virome Analysis & Comparison
Frequency & Richness
heatmap, rarefaction curves, statistical tests
Reference Mapping
bowtie2, samtools, coverage plots in Geneious
Differential Abundance
edgeR, species with fold-difference ≥ 2, volcano plot

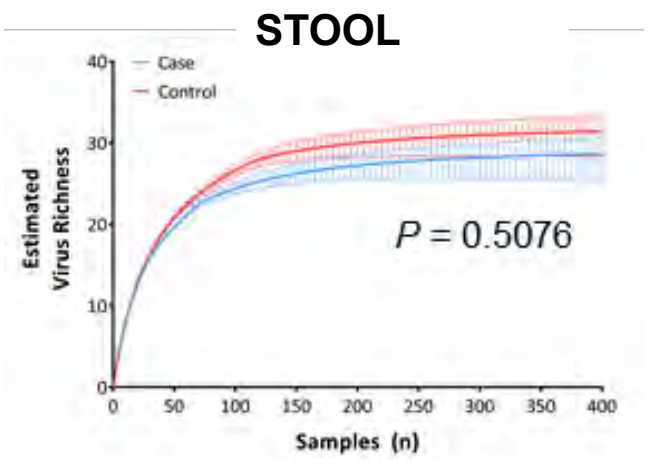
VIGR virome

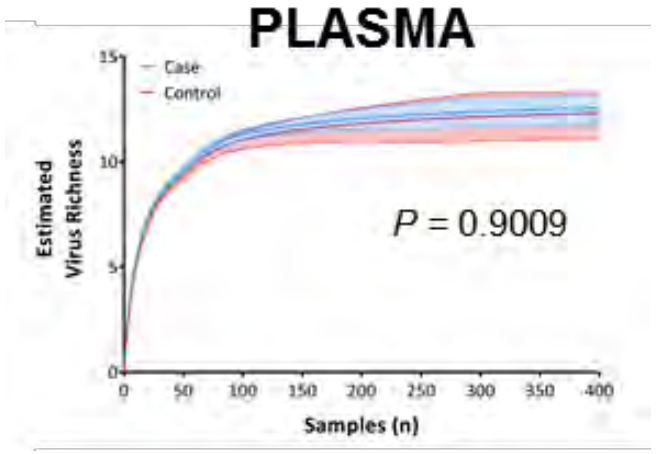


- >2.6 billion raw reads (~14 M filtered reads/sample)
- **28 genera** of vertebrate-infectious viruses
- 75% of stools virus +ve
- 38% of plasma virus +ve
- **62% of children** (58/93) positive for viral nucleic acid

VIGR virome

- No significant difference in viral “richness” at genus level

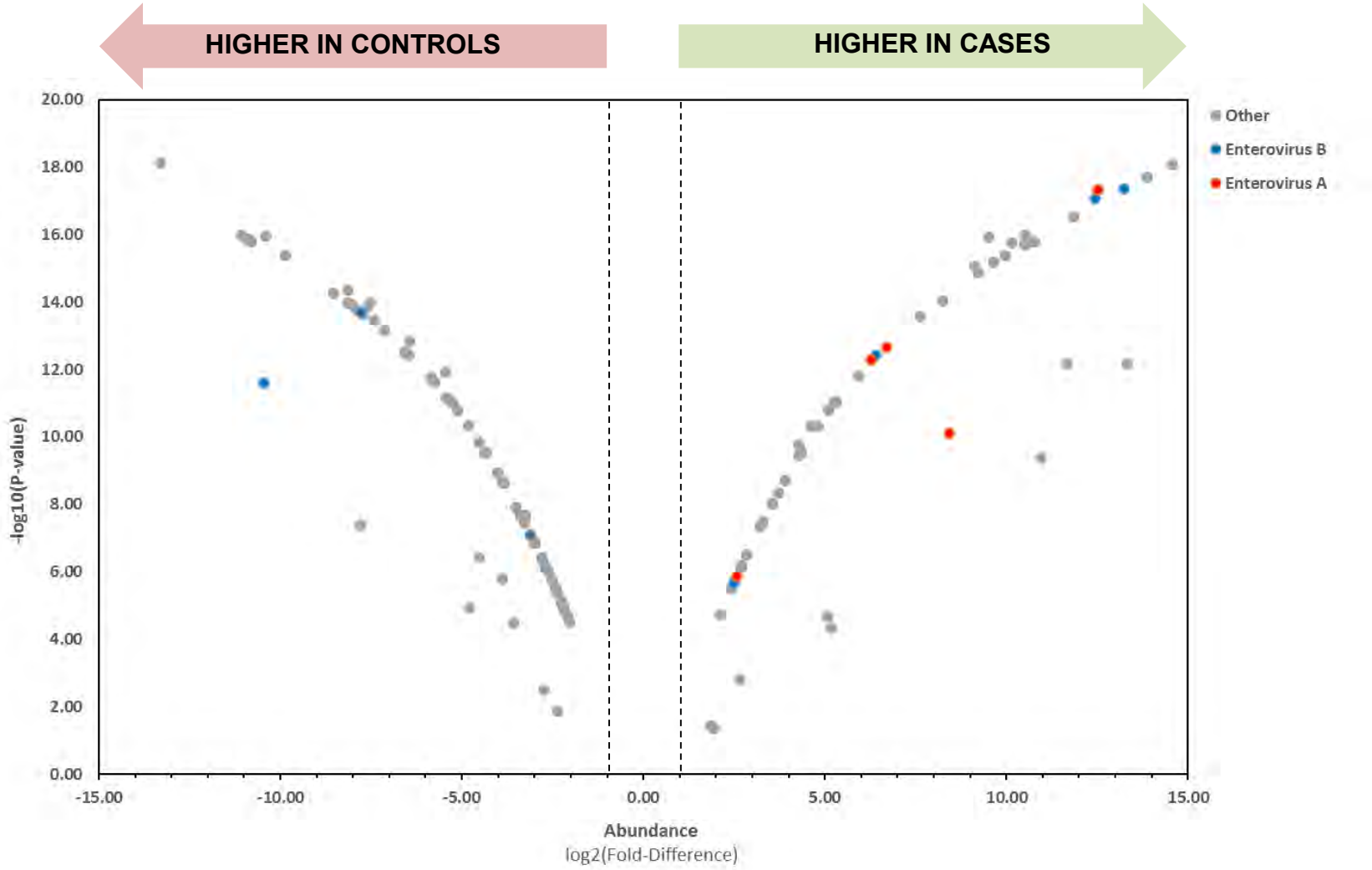




Visit	Case/Control	Viruses detected in Feces	Viruses detected in Plasma	Virus in common
1	Case	None	None	None
2	Case	EV	None	None
3	Case	EV	None	None
4	Case	Mimi_unclassified	None	None
5	Case	None	None	None
6	Case	None	None	None
7	Case	Mamastrovirus Sapovirus	None	None
8	Case	Betatorquevirus	EV Bocaparvovirus Mastadenovirus Roseolovirus	None
9	Case	Anello_unclassified Betatorquevirus	None	None
10	Case	Astro_unclassified Bocaparvovirus Mamastrovirus Picobirnavirus Protoparvovirus	None	None
11	Case	None	None	None
12	Case	None	None	None
13	Case	None	None	None
14	Case	Mastadenovirus	Anello_unclassified Erythroparvovirus	None
15	Case	Picobirnavirus	None	None
16	Case	Picobirnavirus	Betatorquevirus	None
17	Case	Marseille_unclassified	None	None
18	Case	Betatorquevirus	None	None
19	Case	None	None	None
20	Control	Alphatorquevirus Anello_unclassified Bocaparvovirus	None	None
21	Control	Endornavirus EV Mastadenovirus Norovirus Potyvirus Sapovirus	EV	EV
22	Control	None	None	None
23	Control	None	None	None
24	Control	Circovirus	Roseolovirus	None
25	Control	None	None	None
26	Control	Parechovirus	None	None

Sample	Enterovirus	Reference Genome (GenBank Accession)	Reference Sequence Length (nt)	No. Reads Mapped	Reference Coverage (%)
Case					
KWK-258	Coxsackievirus B3	KU574623.1	6,558	56,181	48
KWK-243	Coxsackievirus A2	KX156350.1	7,400	2,354	36
KWK-257	Rhinovirus C	JN815240.1	6,796	249	58
	ECHOvirus E30	EF066392.1	7,334	36,766	54
	Coxsackievirus B3	KR107057.1	7,399	5,293	39
	Coxsackievirus A6	KJ541158.1	7,412	84,400	100
	Coxsackievirus A5	AB114091.1	718	1,770	25
KWK-241	ECHOvirus E18	HM777023.1	7,413	2,555	49
KWK-267	Coxsackievirus A6	AB779616.1	7,434	40,292	25
	Coxsackievirus A2	KC879532.1	885	116,886	100
	Coxsackievirus A14	KP036483.1	7,400	3,720	14
	Coxsackievirus A8	KP765687.1	7,396	39,992	19
Control					
KWK-291	Rhinovirus A	JN798576.1	6,860	3,395	76
KWK-300	Coxsackievirus A6	KJ541158.1	7,412	208	12
KWK-424	Enterovirus A71	HQ647175.1	7,419	4,061	68
	Rhinovirus C	KF688606.1	6,928	136	13
KWK-426	Coxsackievirus B5	KT285015.1	6,937	203,415	42
	ECHOvirus E25	KX139460.1	7,436	2,103	13
KWK-259	Coxsackievirus B4	KX752784.1	7,372	258	20

- 129 differentially abundant viruses (≥2-fold, p<0.05, False Discovery Rate <5%)
- 5 EV-A genotypes more abundant in the gut of cases



How does it relate to other studies?

- Higher rate of virus positivity
  - European DIPP study of 16 IA+ cases **(Kramna et al., *Diabetes Care* 2015)**
  - gut virus positivity of 10.4% (10/96 samples) using 50/100K raw threshold
  - 36% (23/64) positivity in VIGR feces

- Emphasis on identifying statistically significant differences in virus frequency

- Importance to also compare differential abundance of viruses

- Strong focus on EV-B genotypes – warrant further research of EV-As

- Predominance of EV-A consistent with:
  - Natural EV circulation in infants **(Witsø et al., *J Clin Microbiol* 2006)**
  - 10 yr follow-up of 129 children who developed persisten islet autoimmunity in Finland **(Honkanen et al., *Diabetologia* 2017)**

Working model

- Gut as a reservoir for transmission into pancreas
- Anatomically very close, share common lymphatic & vasculature networks
- Sufficient viral load required to establish persistent infection in the gut
- Maintain chronic inflammation milieu that may promote islet autoreactivity by bystander activation
- Preceding intestinal inflammation required for the development of β-cell autoimmunity in mice
- More frequent detection of EV RNA and gut inflammation in T1D patients (small-bowel mucosal biopsy samples of 39 cases with matched controls)

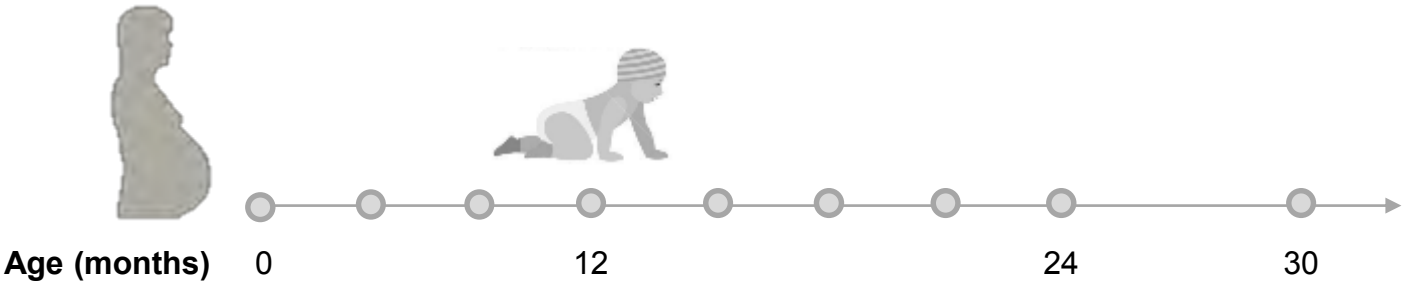
ENDIA environmental determinants of islet autoimmunity

www.endia.org.au

Home About Us What's Involved Regional Program Health Professionals Researcher Resources News Contact

UNSW AUSTRALIA

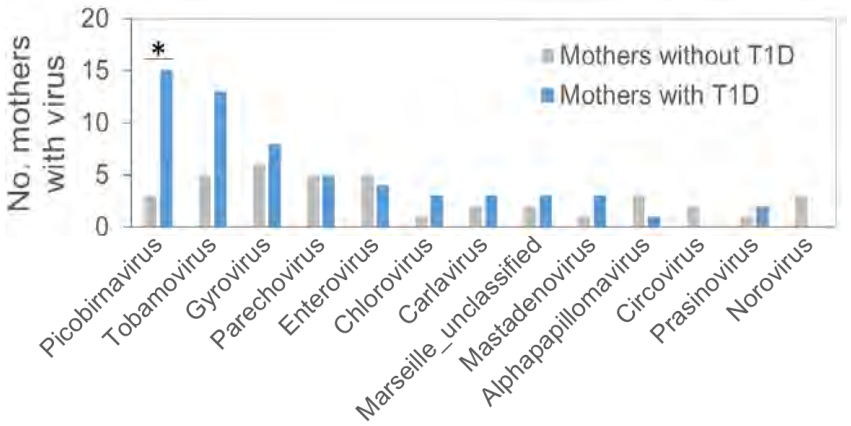
ENDIA virome



- Nationwide prospective T1D birth cohort
- 1,400 at risk infants and mothers
- 3rd year – 6 T1D, 21 with islet autoimmunity
- Pregnancy gut virome: 61 mothers, 1-3 trimesters (n=124)
- Infant gut virome: 25 infants, 4 x visits (n=100)

ENDIA virome

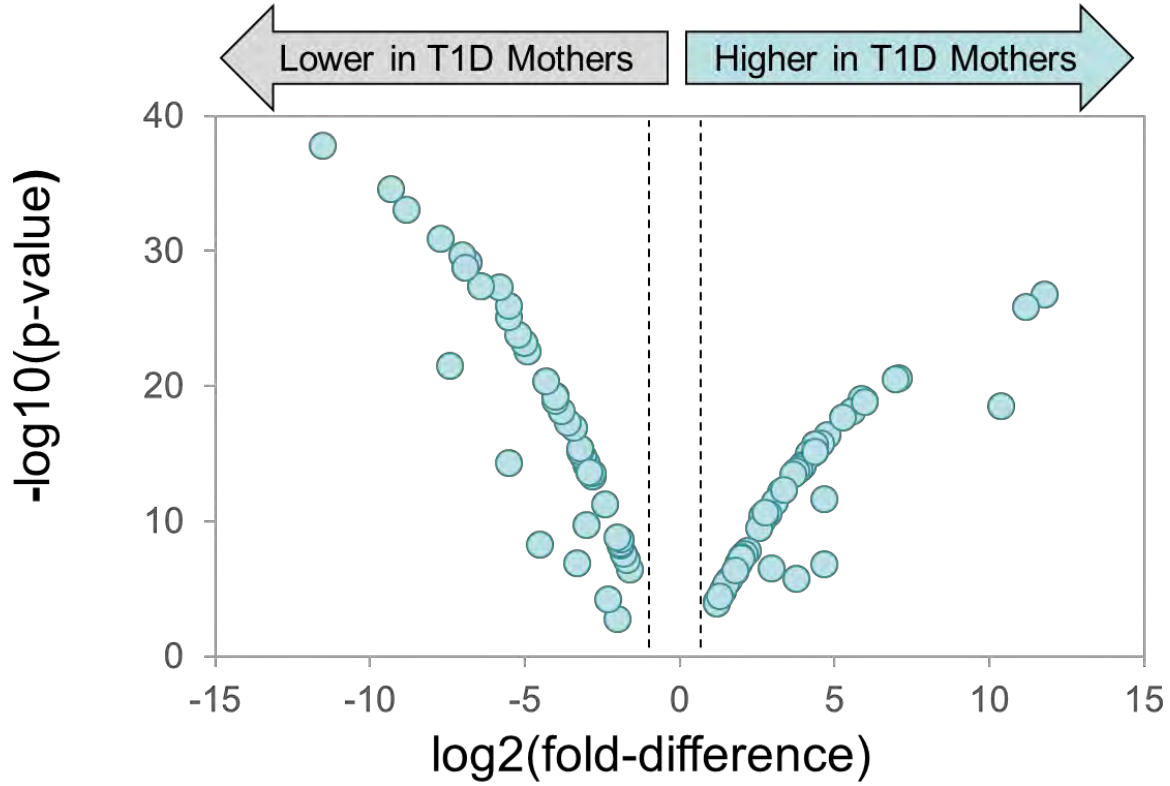
- Gut virome of 61 mothers & 25 infants during pregnancy and early life (n=224 samples)
- >2 billion raw reads (~13M filtered reads/sample)
- 28 genera of vertebrate-infecting viruses in pregnancy
- Trend to higher virus positivity in mothers with T1D vs without (64% vs 50%, p=0.14)
- Picobirnavirus more prevalent in mothers with T1D (p=0.006)



Frequency of viruses detected pregnancy stools (n=124)
**p*=0.006.

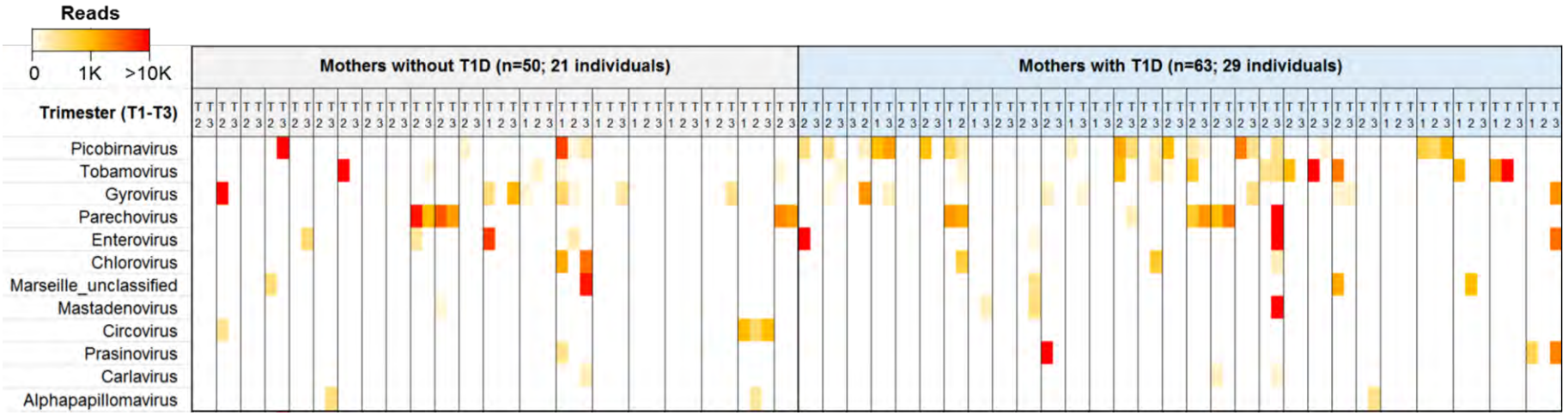
ENDIA virome

- 77 viruses differentially abundant between gut of mothers with vs without T1D

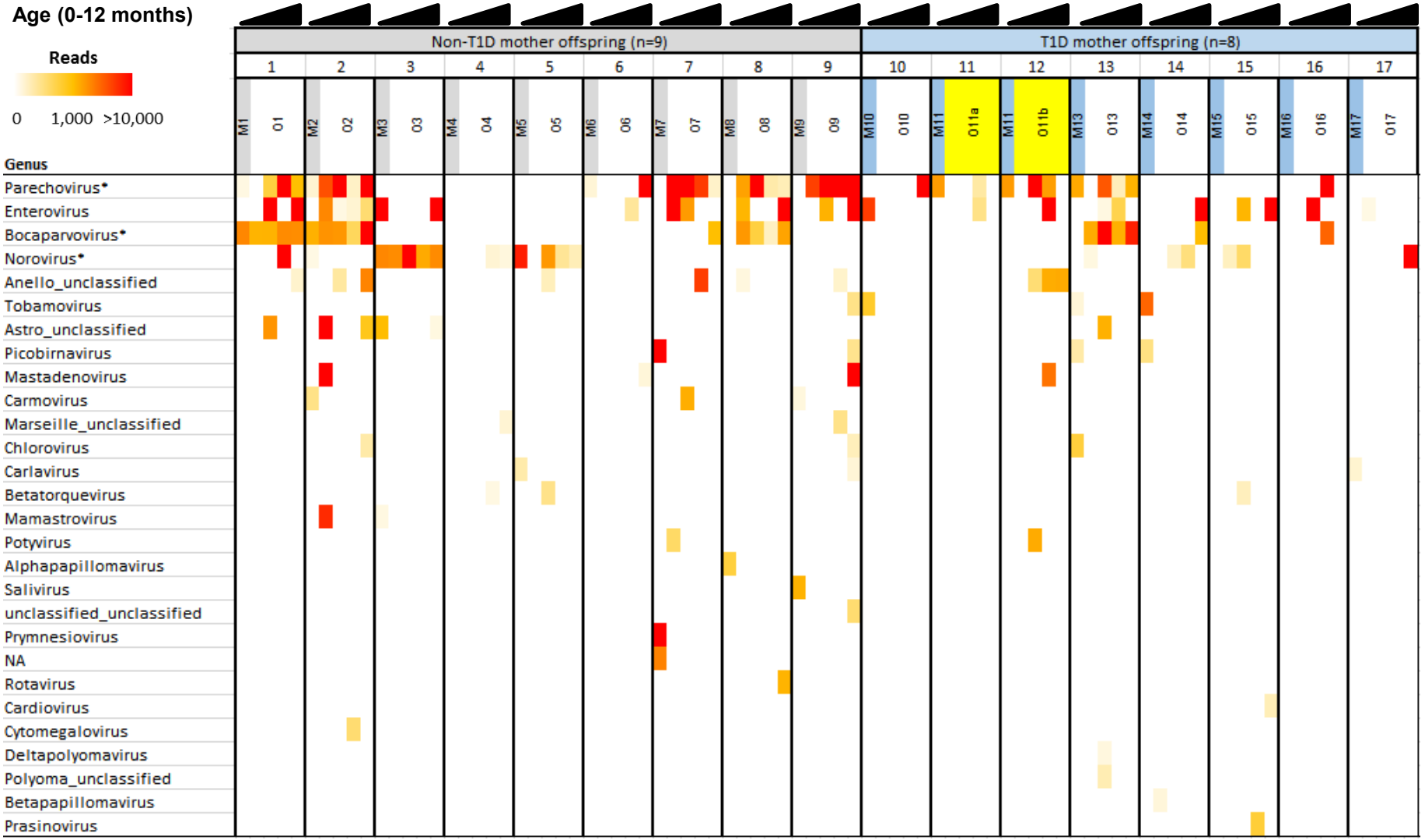


ENDIA virome

- Picobirnavirus, Parechovirus and Circovirus across multiple trimesters
- Persistent/recurring infection?



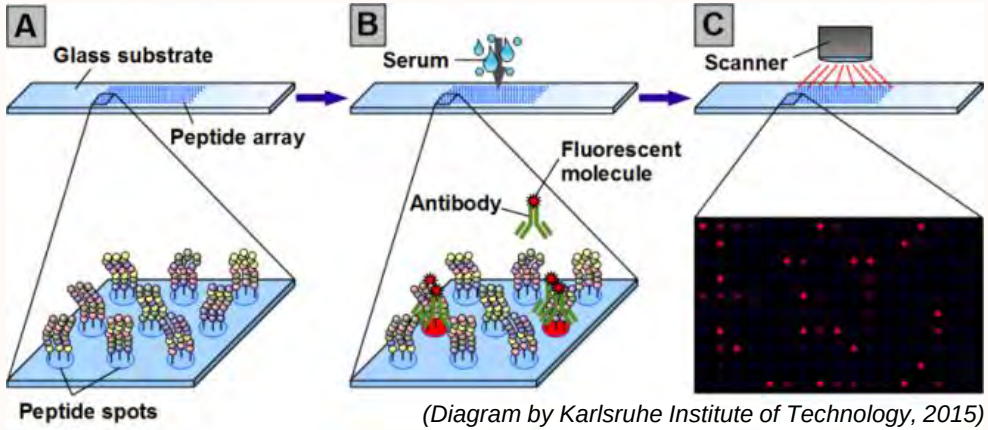
- 24 different genera of vertebrate viruses detected in early life
- 17 mother-infant pairs reveal potential vertical transmission of viruses (*)



Comprehensive serological profiling

High density viral peptide array

12 mer linear epitopes



(Diagram by Karlsruhe Institute of Technology, 2015)

3 million-feature peptide array that accurately displays a repertoire of linear epitopes for all viruses known to infect vertebrates (207 viral taxa)

- Peptides synthesised *in situ* then tiled with 2 amino acid offset
- Enables comprehensive interrogation of a child's humoral immune response to all past virus exposures using <100 µL of serum

- **Virology Research Laboratory:**
  - Prof. Maria Craig.
  - Prof. William Rawlinson
  - Jessica Horton (Hons)
  - Sonia Isaacs (PhD)
  - Digby Allen (Med Hons)
  - Nurses: Jacki, Kelly, Deb
- **UNSW:**
  - Prof. Marc Wilkins
  - Dr. Ignatius Pang
  - Dr. Rowena Bull
  - Dr. Fabio Luciani
- **Center for Infection & Immunity:**
  - Prof. Ian Lipkin
  - A/Prof. Thomas Briese
  - Komal Jain (Bioinformatics)
  - Nishit Bhuva (Sequence-capture)
- **ENDIA Study:**
  - Prof. Jennifer Couper (CI)
  - All investigators
  - Study coordinators
  - Study team
- **Mothers & Children supporting ENDIA**

ENDIA

Serology & virology division
diagnosis.research.teaching

UNSW AUSTRALIA

SEALS excellence in diagnosis research teaching

UNSW AUSTRALIA

Rebecca L. Cooper
Medical Research Foundation

Australian Type 1 Diabetes
Clinical Research Network

diabetes
australia

JDRF IMPROVING LIVES. CURING TYPE 1 DIABETES.

The Sydney Partnership for Health, Education, Research & Enterprise

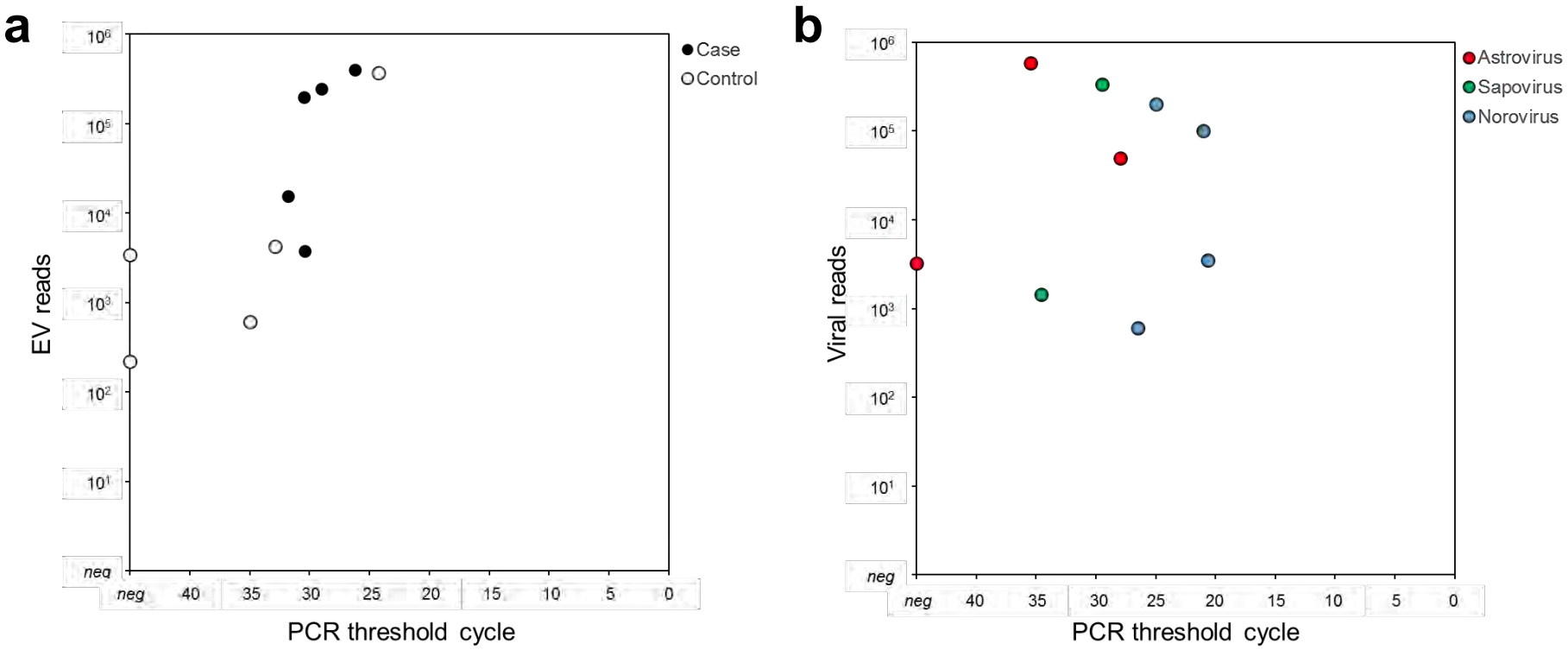
Maridulu Budyari Gumal

UNSW AUSTRALIA

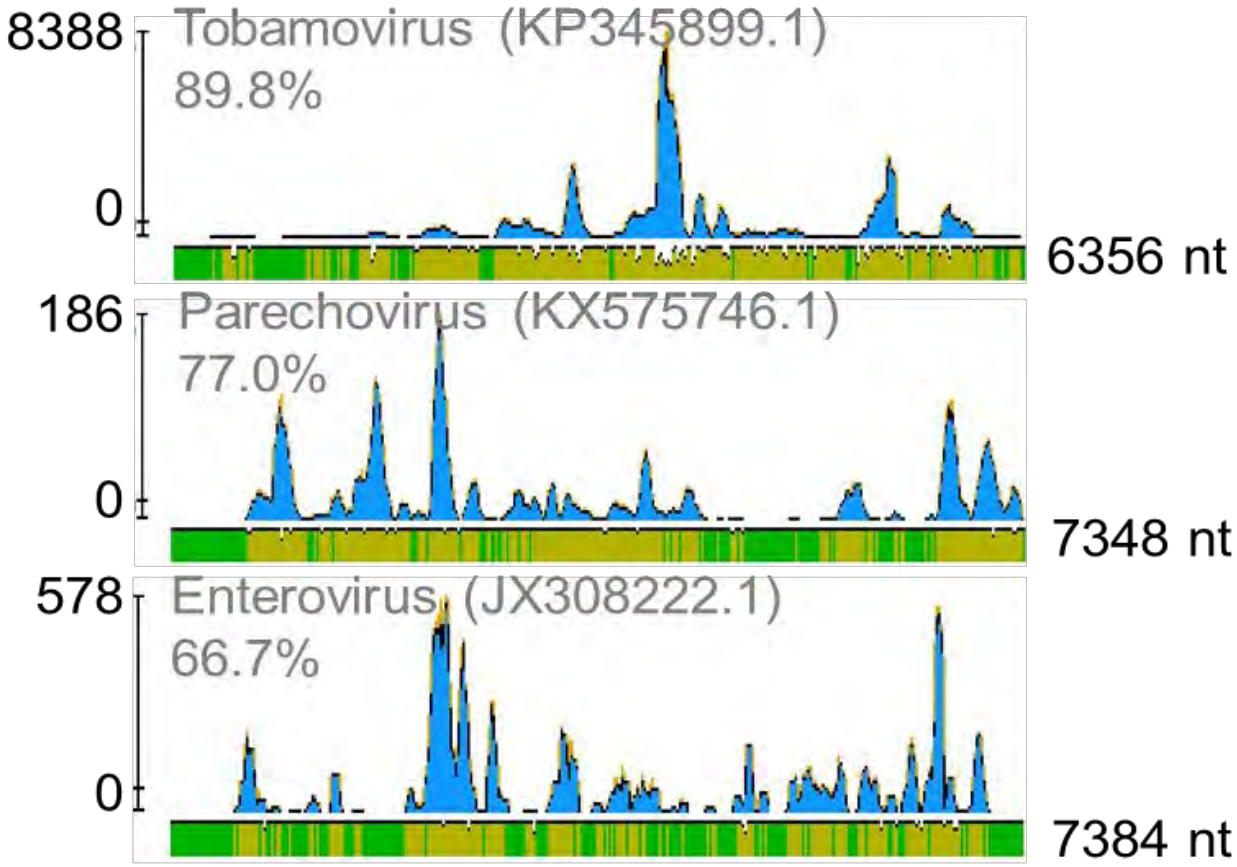

EDWIN STREET
KITCHENNET
Ph: 9181 3888
KitchenNet.com.au
THE STOOL SHOP
STOOL SHOP
147b
9719
BARS·STOOLS
OILS

Supporting Slides

Virus	Main Finding	Reference
Rubella	Diabetes in patients within a congenital rubella cohort	[44]
	Majority of patients with diabetes from [44] classified as having T2D	[33]
	No islet autoantibodies detected in congenital rubella syndrome cohort	[34]
	Epidemiological study showed association between rubella and T1D	[37]
Mumps	Infection during mumps epidemic more common in children who subsequently developed T1D	[36]
	Epidemiological study showed association between mumps and T1D	[37]
CMV	CMV detected more often in patients with T1D than controls	[38]
	Prospective cohort study found no association between CMV infection in infancy and T1D	[40]
	No association between perinatal CMV infection and T1D	[39]
Rotavirus	Majority of children at risk for T1D developed islet autoantibodies following rotavirus seroconversion	[41]
	No association between rotavirus infection and T1D in children	[42]
	No association between rotavirus infection and islet autoimmunity or T1D in children	[43]
EV	Higher levels of CVB antibodies in recently diagnosed diabetes patients	[45]
	Systematic review of 26 serological studies, inconclusive evidence for a role of CVB in T1D	[46]
	CVB1 antibodies more common in children with T1D	[47]
	VP1 present in pancreas samples from patients with T1D	[48, 49]
	CVB4 infection in islets of 3 of 6 patients with T1D, associated with impaired β cell function	[50]
	VP1 colocalises with viral response element PKR in insulin-containing islet cells in patients with T1D	[51]
	Weak association between EV in blood and islet autoantibody detection	[52]
	No association between EV in stool and T1D	[53]
	Significant association between EV infection and T1D, particularly with severe ketoacidosis	[54, 55]
	Meta-analysis of molecular studies determines >10-fold higher rate of EV in infection in T1D compared to controls	[56]



Comparison of virus detection in VirCapSeq-VERT versus targeted real-time PCR. (a) Five case and five control feces identified as EV-positive by VirCapSeq-VERT. Two of the control specimens tested negative (*neg*) by EV-specific qPCR. (b) Multiplex qPCR validation of fecal specimens identified as positive for astroviruses, sapoviruses and noroviruses by VirCapSeq-VERT. Astrovirus positivity could not be qPCR validated in one specimen.



High virus genome coverage achieved using VirCapSeq-VERT. Filtered reads were mapped to reference genome accessions identified by BLAST (megablast). Partial/complete genome sequences were obtained from GenBank, alignments performed using bowtie2 and coverage plots generated in Geneious R9.