wellcome centre human genetics

What is Viromics and why does it matter?

Oxford Viromics Overview

HiDi Day, 1 October 2018

Rory Bowden – WHG-OGC, OxSingleCell, Oxford Viromics

What is Viromics?

Viromics is just the study of viruses at genome and population scale.

The study of viruses becomes vastly more tractable once we get **complete genome information**.

Modern genomics techniques give us a chance to access **whole virus genomes cheaply, completely and at scale**.

Host factors in virus infection only relevant once we **fully define the virus**.

Viromics is intimately interconnected with **Immunology**

Viromics can be used in **diagnosis, tracing, monitoring, treatment, epidemiology, pathogenesis, …**

What is Oxford Viromics?

Oxford Viromics is…

- **An Idea**

I. Make available sequencing and analysis for human virus-containing samples for small and large studies.

II. Proof of concept of the value of virus detection and genomics in clinical contexts.

III. Enable focused funding applications.

IV. A pipeline for virus detection and genetic characterization to enable clinical trials that define the role of state-of-the-art sequencing technology in clinical virology.

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- **Some People**

Paul Klenerman (TGU/Medawar)

David Bonsall (Medawar/WHG/BDI)

Rory Bowden (WHG)

Tanya Golubchik (WHG/BDI)

Mariateresa de Cesare (WHG)

Azim Ansari (WHG/Medawar)

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- **Collaborations**

Barnes – Medawar/StopHCV

Matthews – Medawar/BRC

Pybus – Zoology

Simmonds – Medawar/BRC

Crook – MMM/BRC

Fraser – BDI

Maiden – Zoology

Lythgoe - BDI

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- Collaborations
- **Some Funding**

Oxford Wellcome Trust ISSF + WHG

Oct 2016 – Sept 2018

2 half-time posts for 2 years

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- Collaborations
- Some Funding
- **Some Technology**

Peter Medawar Building:

- extraction platform

Oxford Genomics Centre:

- liquid handling
- quantitation and QC
- sequencing: Illumina (x4), Nanopore

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- Collaborations
- Some Funding
- Some Technology
- **Some Lab Methods**

Optimising available methods:

- Extraction platform
- RNA-seq kits
- Sequence enrichment kits
- RNA-and-DNA libraries
- Streamlining, miniaturization, automation
- Nanopore

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- Collaborations
- Some Funding
- Some Technology
- Some Lab Methods
- **Some Analytical Methods**

An initial pipeline …

- `snork`
- metagenomics pipeline(s)
- `shiver` and `phyloscanner`
- nanopore-specific tools
- probe-design tools

What is Oxford Viromics?

Oxford Viromics is…

- An Idea
- Some People
- Collaborations
- Some Funding
- Some Technology
- Some Lab Methods
- Some Analytical Methods
- **A Symposium**

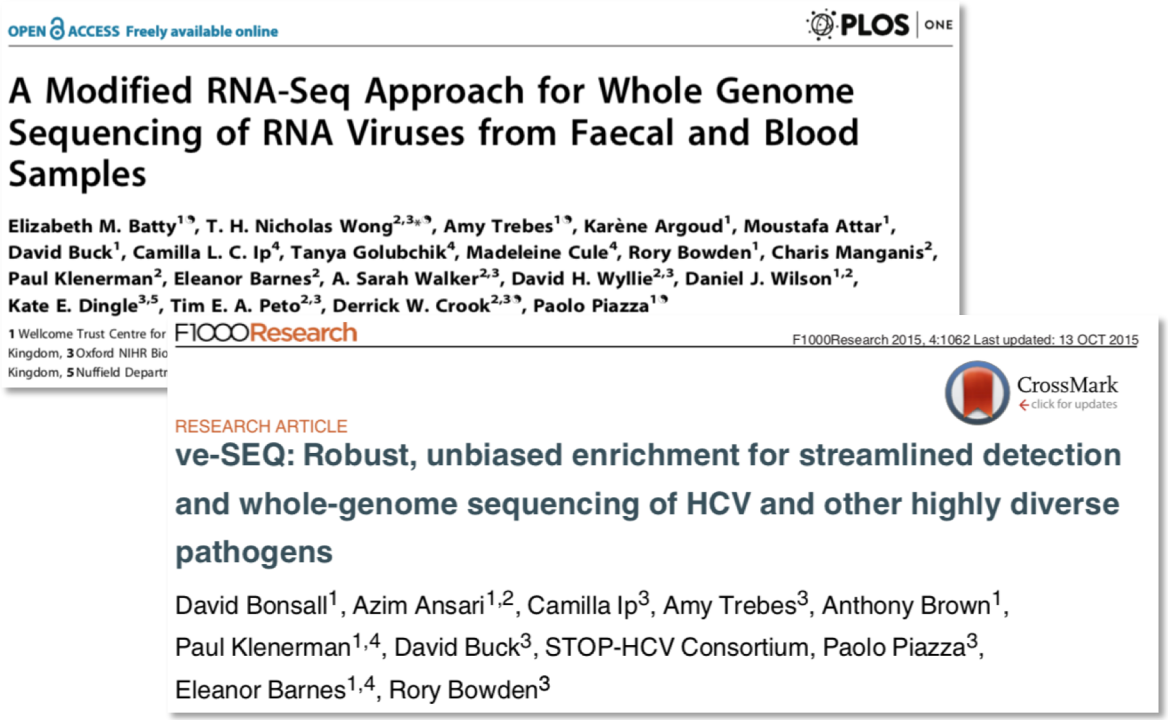
Save the Date:

11 January 2019

James Martin School

Proof-of-concept studies

How we knew Viromics was tractable and worthwhile



OPEN ACCESS Freely available online

PLOS ONE

A Modified RNA-Seq Approach for Whole Genome Sequencing of RNA Viruses from Faecal and Blood Samples

Elizabeth M. Batty[1], T. H. Nicholas Wong[2,3*], Amy Trebes[1], Karène Argoud[1], Moustafa Attar[1], David Buck[1], Camilla L. C. Ip[4], Tanya Golubchik[4], Madeleine Cule[4], Rory Bowden[1], Charis Manganis[2], Paul Klenerman[2], Eleanor Barnes[2], A. Sarah Walker[2,3], David H. Wyllie[2,3], Daniel J. Wilson[1,2], Kate E. Dingle[3,5], Tim E. A. Peto[2,3], Derrick W. Crook[2,3], Paolo Piazza[1]

1 Wellcome Trust Centre for
Kingdom, 3 Oxford NIHR Bio
Kingdom, 5 Nuffield Departr

F1000Research

F1000Research 2015, 4:1062 Last updated: 13 OCT 2015

CrossMark click for updates

RESEARCH ARTICLE

ve-SEQ: Robust, unbiased enrichment for streamlined detection and whole-genome sequencing of HCV and other highly diverse pathogens

David Bonsall[1], 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]

- RNA-seq has advantages over PCR-seq
- Probe enrichment is robust and predictable
- Can design probe sets to any arbitrary group of pathogens: 10kb – 5Mb++

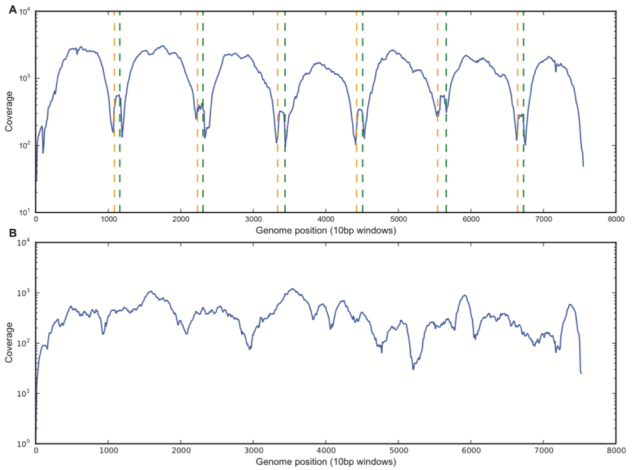


Figure 1. Coverage profiles of one Norovirus sample from amplicon and direct RNA sequencing. A – Coverage across the genome for one Norovirus sample sequenced from PCR amplicons (others similar). Green and orange dotted lined mark the locations of the PCR primers used to generate the amplicons. B – coverage across the genome for the same Norovirus sample sequenced directly from RNA.
doi:10.1371/journal.pone.0066129.g001

Norovirus

PCR

RNA-seq

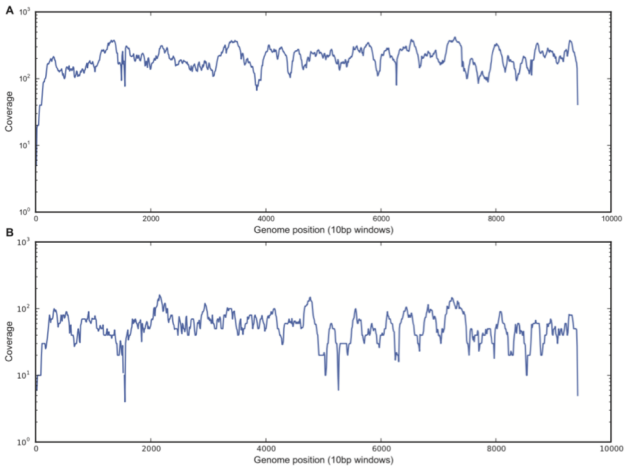


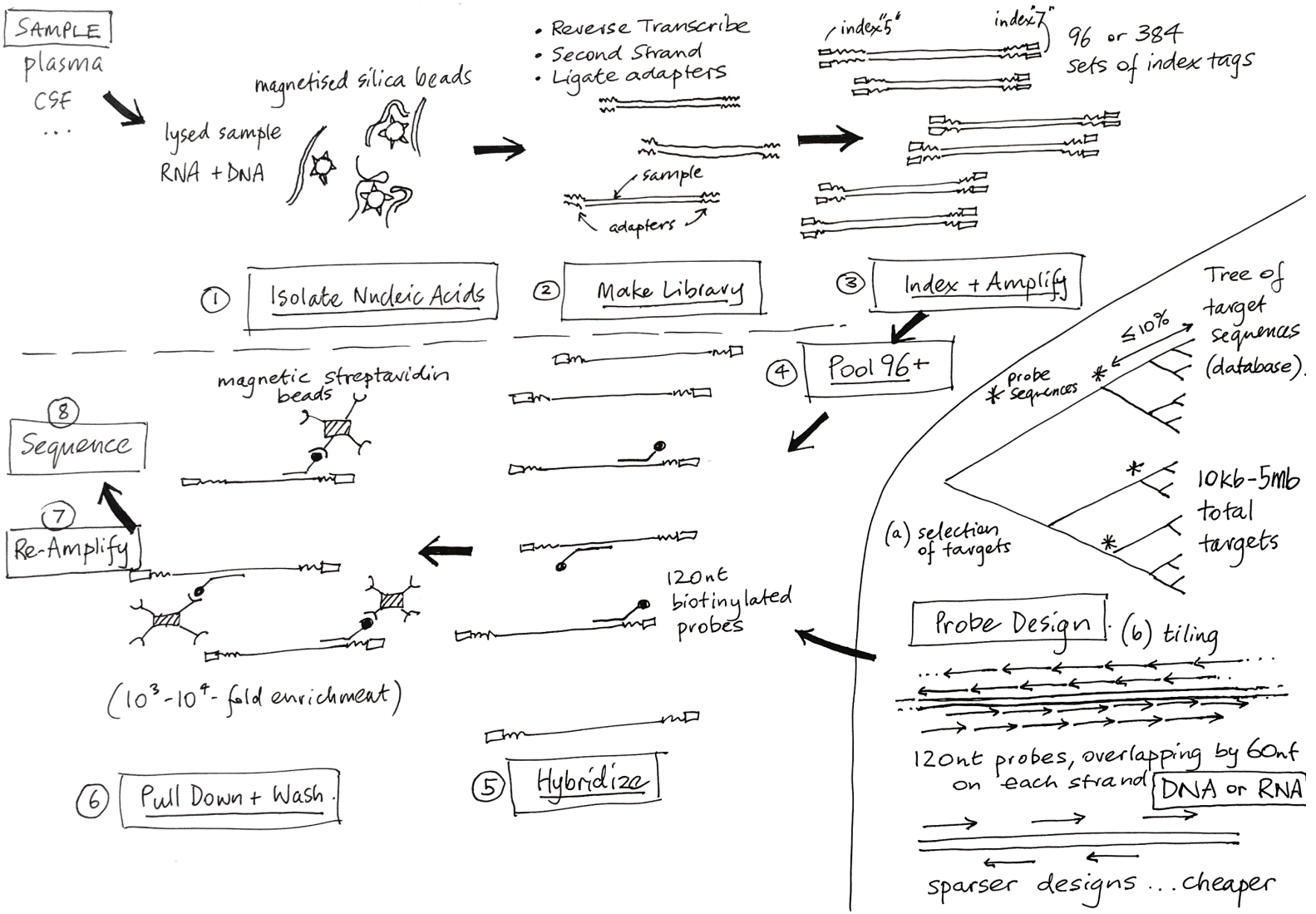
Figure 2. Coverage across the genome for two Hepatitis C samples sequenced directly from RNA.
doi:10.1371/journal.pone.0066129.g002

Hepatitis C virus

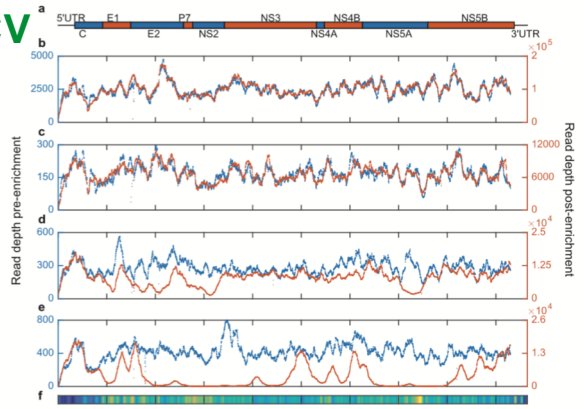
RNA-seq vs PCR

RNA-seq:

- is less sensitive than PCR
- doesn't depend on matching primers
- can work for degraded samples
- needs fewer tubes


SAMPLE
plasma
CSF
...
lysed sample
RNA + DNA
magnetised silica beads
• Reverse Transcribe
• Second Strand
• Ligate adapters
sample
adapters
index "5"
index "7"
96 or 384
sets of index tags
① Isolate Nucleic Acids
② Make Library
③ Index + Amplify
④ Pool 96+
Tree of target sequences (database).
≤10%
probe sequences
10kb-5Mb total targets
(a) selection of targets
Probe Design
(b) tiling
120nt probes, overlapping by 60nt on each strand
DNA or RNA
sparser designs ... cheaper
120nt biotinylated probes
⑤ Hybridize
magnetic streptavidin beads
⑥ Pull Down + Wash
⑦ Re-Amplify
⑧ Sequence
(10³-10⁴- fold enrichment)

HCV



← perfect match

← same subtype

← same genotype

← different genotype

red = enriched blue = metagenomic

Figure 2. Enrichment efficiency decreases with phylogenetic distance. Read depth across the genome before (blue, left axis) and after (red, right axis) enrichment with a single-sequence subtype 1a probe set. **a**. The HCV genome comprises 5′ and 3′ untranslated regions (UTRs) and a large central segment encoding a single polyprotein that is cleaved into ten proteins. **b**. A subtype 1a sample enriched with probes derived from its own consensus sequence yields coverage patterns across the genome essentially identical to metagenomic sequencing. **c**. A distinct subtype 1a sample produces highly similar but non-identical patterns of pre- and post-enrichment genomic coverage. **d**. A subtype 1b sample yields low read depths at loci that are relatively divergent from the 1a probe sequence (E1, E2, NS2 and NS5a). **e**. Sequence capture of a sample from a different genotype, 3a, is poor across large segments of the genome. **f**. Heat map representing average diversity (calculated as Shannon entropy) among 165 HCV reference genomes. Nucleotide diversity varies dramatically across the genome and tracks drops in enrichment efficiency between phylogenetically distinct probe-target combinations.

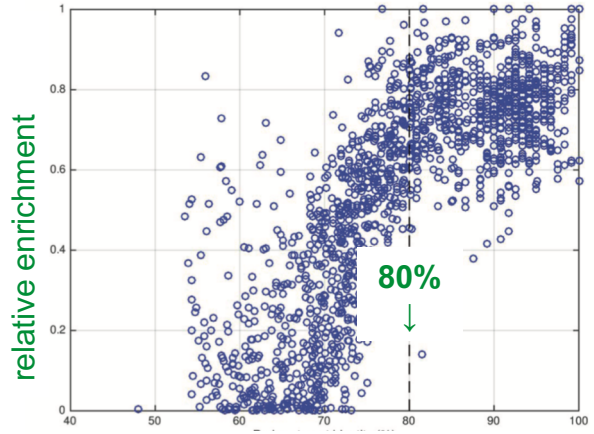


Figure 3. Enrichment efficiency is direct L was sequenced before and after enrichment with a single-ge[illegible] was plotted against the maximum identity between target and any probe. Read depth ratio was normalized by giving the most efficiently enriched probe position (in the highly conserved 5′ UTR) a value of 1. Maximal enrichment is observed where probe-target identity exceeds approximately 80% and enrichment decreases dramatically as identity falls below 80%.

Enrichment Sequencing

- Like exome sequencing
- Biotinylated DNA or RNA baits
- Enrich pooled libraries
- Probe-based enrichment is robust and predictable, even for HCV
- Typically 10^3 – 10^4-fold enrichment
- Tolerant of substantial sequence divergence
- Can design probe sets to any arbitrary group of pathogens: 10kb–5Mb+

nature genetics

Genome-to-genome analysis highlights the effect of the human innate and adaptive immune systems on the hepatitis C virus

M Azim Ansari[1–3,11], Vincent Pedergnana[1,11], Camilla L C Ip[1,3], Andrea Magri[3], Annette Von Delft[3], David Bonsall[3], Nimisha Chaturvedi[4], Istvan Bartha[4], David Smith[3], George Nicholson[5], Gilean McVean[1,6], Amy Trebes[1], Paolo Piazza[1], Jacques Fellay[4], Graham Cooke[7], Graham R Foster[8], STOP-HCV Consortium[9], Emma Hudson[3], John McLauchlan[10], Peter Simmonds[3], Rory Bowden[1], Paul Klenerman[3], Eleanor Barnes[3], Chris C A Spencer[1]

Outcomes of hepatitis C virus (HCV) infection and treatment depend on viral and host genetic factors. Here we use human genome-wide genotyping arrays and new whole-genome HCV viral sequencing technologies to perform a systematic genome-to-genome study of 542 individuals who were chronically infected with HCV, predominantly genotype 3. We show that both alleles of genes encoding human leukocyte antigen molecules and genes encoding components of the interferon lambda innate immune system drive viral polymorphism. Additionally, we show that *IFNL4* genotypes determine HCV viral load through a mechanism dependent on a specific amino acid residue in the HCV NS5A protein. These findings highlight the interplay between the innate immune system and the viral genome in HCV control.

HCV infection presents a major health burden, with more than 185 million people being infected worldwide[1], which can lead to liver failure and hepatocellular cancer in infected individuals. Genetic variations in both the host and the virus are associated with important clinical outcomes. Genetic polymorphisms in the host, most notably in the interferon (IFN) lambda 3 (*IFNL3*) and *IFNL4* loci, are associated with spontaneous clearance of the virus, response to treatment, viral load and progression of liver disease[2–6]. Viral genotypes and distinct viral genetic motifs have been associated with the response to interferon-based therapies[7,8], whereas resistance-associated substitutions (RASs) have been identified for most of the new oral direct-acting antiviral (DAA) drugs[9–12]. HCV can be divided into seven major genotypes, and most of the genetic data acquired to date has focused on HCV genotype 1, with a lack of data for other genotypes. HCV genotype 3 is of particular interest, as this genotype is known to infect 53 million people globally[13] and is associated with a higher failure rate to DAA therapies[14,15].

Previous work, including candidate gene studies of the association between the human leukocyte antigen (HLA) type I proteins and the HCV genome[16,17], has shown that within-host virus diversity evolves in response to the host adaptive immune system. HLA molecules are expressed on most cell types, and they present viral peptides (epitopes) to cytotoxic T lymphocytes (CTLs), which kill infected cells. CTL-mediated killing of virus-infected cells drives the selection of viral polymorphisms ('escape' mutations) that abrogate T cell recognition[18]. Understanding how host HLA molecules affect viral selection has important implications for the development of HCV-specific T cell vaccines that aim to prevent infection[19,20]. A comprehensive host genome to viral genome analysis at scale will assess the relative contribution of host HLA molecules in driving changes in the HCV genome, and it might also identify other host genes that have a key role in shaping the HCV genome.

We generated data from a cohort of 601 HCV-infected patients (from the BOSON[21] clinical trial) to systematically look for associations between host and virus genomes, exploiting the fact that while the host genome remains fixed the virus mutates, allowing it to evolve during infection. For this, we developed a targeted viral enrichment methodology[22,23] to obtain whole HCV genomes, and we used high-throughput genotyping arrays in combination with statistical imputation to obtain data for nucleotide polymorphisms across the human genome and the alleles of genes encoding HLA molecules[24,25] (hereafter referred to as HLA genes). We provide evidence that polymorphisms relevant to the innate (*IFNL4*) and adaptive immune systems (HLA genes) are associated with HCV sequence polymorphisms. We show that an interaction between host *IFNL4* genotypes and an amino acid residue in the HCV NS5A protein determines HCV viral load. By assessing viral evolution in individuals with different *IFNL4*

[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [2]Oxford Martin School, University of Oxford, Oxford, UK. [3]Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine and the NIHR Oxford BRC, University of Oxford, Oxford, UK. [4]School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [5]Department of Statistics, University of Oxford, Oxford, UK. [6]Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK. [7]Wright-Fleming Institute, Imperial College London, London, UK. [8]Queen Mary University of London, London, UK. [9]A full list of members and affiliations appears at the end of the paper. [10]MRC–University of Glasgow Centre for Virus Research, Glasgow, UK. [11]These authors contributed equally to this work. Correspondence should be addressed to C.C.A.S. (chris.spencer@well.ox.ac.uk) or E.B. (ellie.barnes@ndm.ox.ac.uk).

Received 9 August 2016; accepted 10 March 2017; published online 10 April 2017; doi:10.1038/ng.3835

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

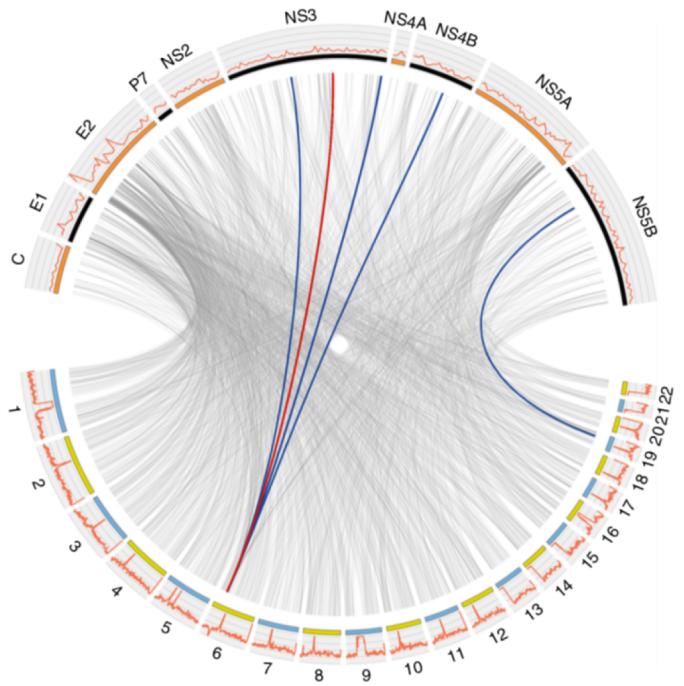
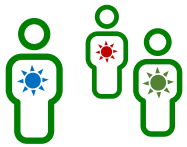
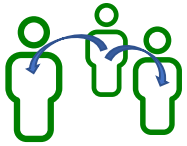


Figure 1 Human-to-HCV genome-wide association study in 542 patients. The lower arc shows the human autosomes from chromosomes 1 to 22, and the upper arc shows the HCV proteome from the core protein (C) to NS5B. The red line represents the most significant association ($P < 2 \times 10^{-11}$). The four blue lines represent suggestive associations ($P < 4 \times 10^{-9}$). The thin gray lines represent associations with $P < 10^{-5}$. The outer mini-panels represent, on the upper arc, the viral diversity as measured by Shannon entropy and, on the lower arc, the density of human SNPs in bins of 1 Mb, with higher values further away from the center for both the upper and lower arcs.

Viromics as a complete solution

for management of a chronic virus infection

Measurement:	Viral load	Genotype	Transmission network	Drug resistance levels
				
Sequencing design:	Quantitative standards	Unbiased probe capture	Minimal PCR, Fragment size selection	Quantitative sequencing, Optimisation for low viral loads
Analysis design:	PCR duplicate removal	Accurate mapping, Consensus calling	Ancestral host-state reconstruction	Haplotype calling, HIVdb

HIV: BDI – Christophe Fraser

BIG DATA INSTITUTE

HPTN
HIV Prevention
Trials Network

PopART
HPTN 071

HIV genotyping and phylogenetics in HPTN 071 (PopART)

HIGH-THROUGHPUT SEQUENCING TO ASSESS VIRAL LOAD, GENOTYPE, DRUG RESISTANCE AND TRANSMISSION

PEPFAR

3ie International Initiative for Impact Evaluation

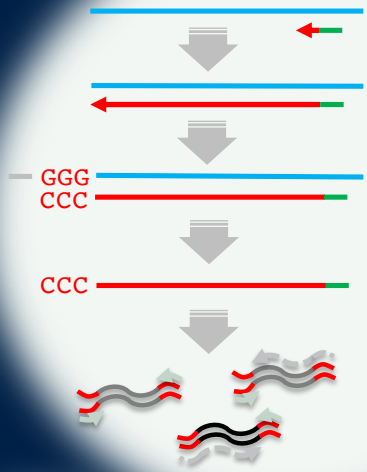
BILL & MELINDA GATES foundation

U.S. NATIONAL INSTITUTES OF HEALTH:
National Institute of Allergy and Infectious Diseases
National Institute of Mental Health
National Institute on Drug Abuse

Nucleic Acids Extraction
EasyMag (Biomerieux)

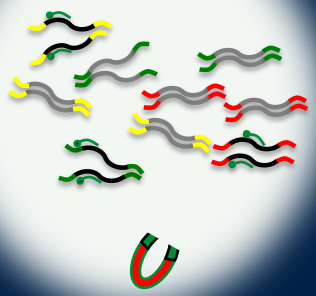
Library prep
SMARTer (TakaraBio) Adapter attachment without ligation

Liquid Handling
Labcyte Echo 525
Automated
384 - well format



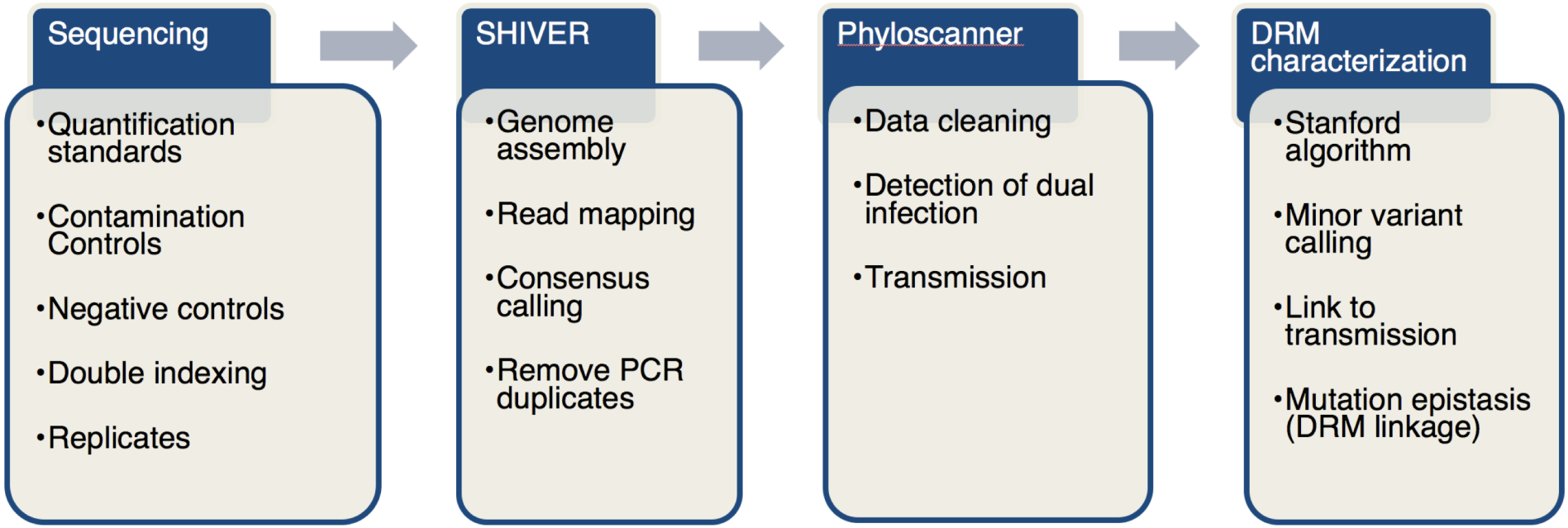
384 samples / week
£30 / sample

Minimal PCR
+ Size selection of larger fragments



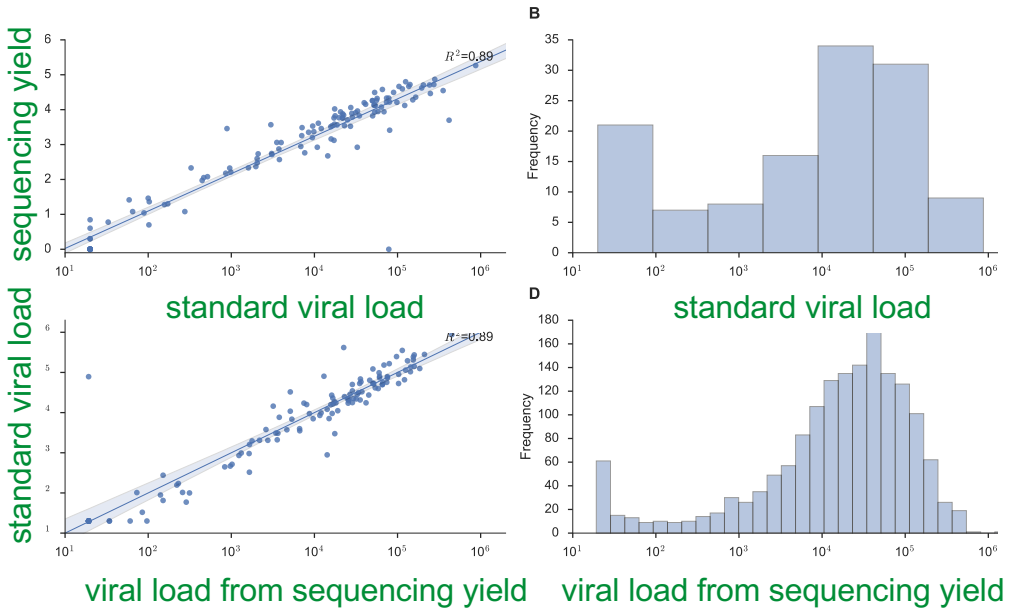
Enrichment
650 Custom DNA oligos to capture full HIV A-D diversity

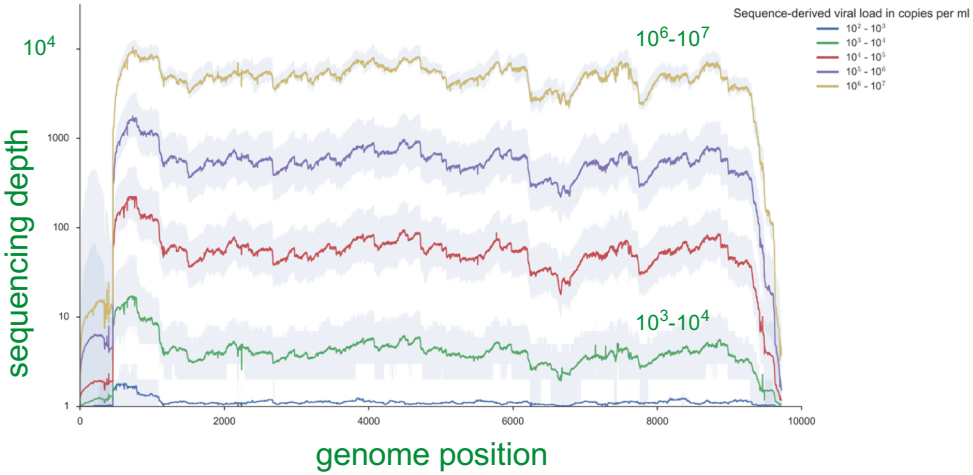
Sequencing
Illumina MiSeq / HiSeq 2500 Rapid
300b / 250b paired reads

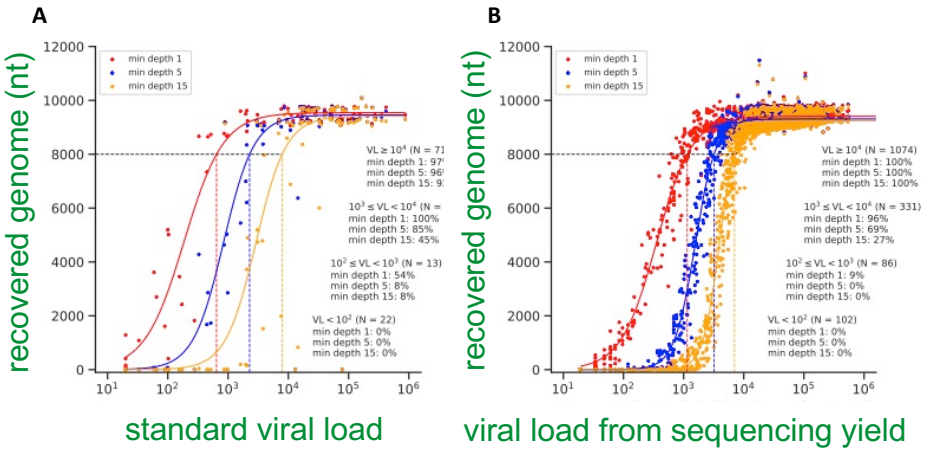

Sequencing
•Quantification standards
•Contamination Controls
•Negative controls
•Double indexing
•Replicates
SHIVER
•Genome assembly
•Read mapping
•Consensus calling
•Remove PCR duplicates
Phyloscanner
•Data cleaning
•Detection of dual infection
•Transmission
DRM characterization
•Stanford algorithm
•Minor variant calling
•Link to transmission
•Mutation epistasis (DRM linkage)

HIV Sequencing:

- Viral load
- Genome coverage
- Sensitivity







A comprehensive genomics solution for HIV surveillance and clinical monitoring in a global health setting

David Bonsall[a,b,*], Tanya Golubchik[a,b,*], Mariateresa de Cesare[b], Mohammed Limbada[c,d], Barry Kosloff[c,d], George MacIntyre-Cockett[b,a], Matthew Hall[a], Chris Wymant[a], M Azim Ansari[b,e], Lucie Abeler-Dörner[a], Ab Schaap[c,d], Anthony Brown[e], Eleanor Barnes[e], Estelle Piwowar-Manning[f], Ethan Wilson[g], Lynda Emel[g], Richard Hayes[d], Sarah Fidler[h], Helen Ayles[c,d], Rory Bowden[b], Christophe Fraser[a]

[a] Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[b] Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom
[c] ZAMBART, University of Zambia, Lusaka, Zambia
[d] London School of Hygiene and Tropical Medicine, London, United Kingdom
[e] Peter Medawar Building for Pathogen Research, University of Oxford, Oxford, United Kingdom
[f] HIV Prevention Trials Network (HPTN) Laboratory Core, Johns Hopkins University, Baltimore, Maryland, USA
[g] Statistical Centre for HIV/AIDS Research, Fred Hutchinson Cancer Research Centre, Seattle, Washington, USA
[h] Department of Medicine, Imperial College London, London, United Kingdom

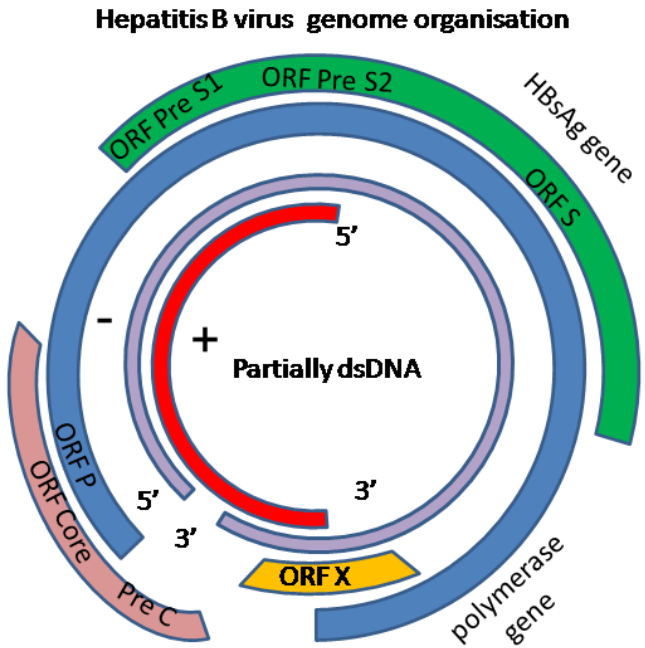
https://www.biorxiv.org/content/early/2018/08/28/397083

HIV Sequencing:

- Clinical management
- Transmissions
- Dynamics of infection
- Drug resistance

Viromics with new methods

Not every virus is ssRNA or dsDNA



https://commons.wikimedia.org/wiki/File:HBV_genome.png

Hepatitis B virus has a circular, partially double-stranded, DNA, virion genome that is not covalently closed.

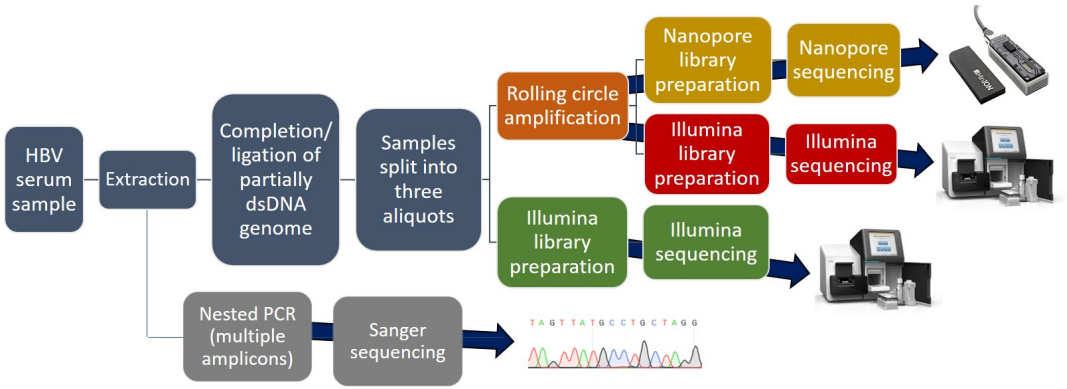
⇢ "completion/ligation" plus "rolling-circle" amplification with HBV primers

= concatemers (chains) of successive full-length genomes

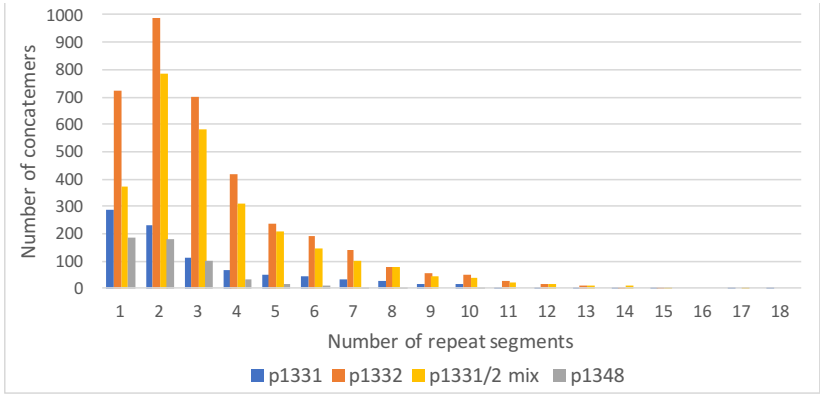
⇢ HBV-enriched; suitable for Illumina/**Nanopore** sequencing, further enrichment.

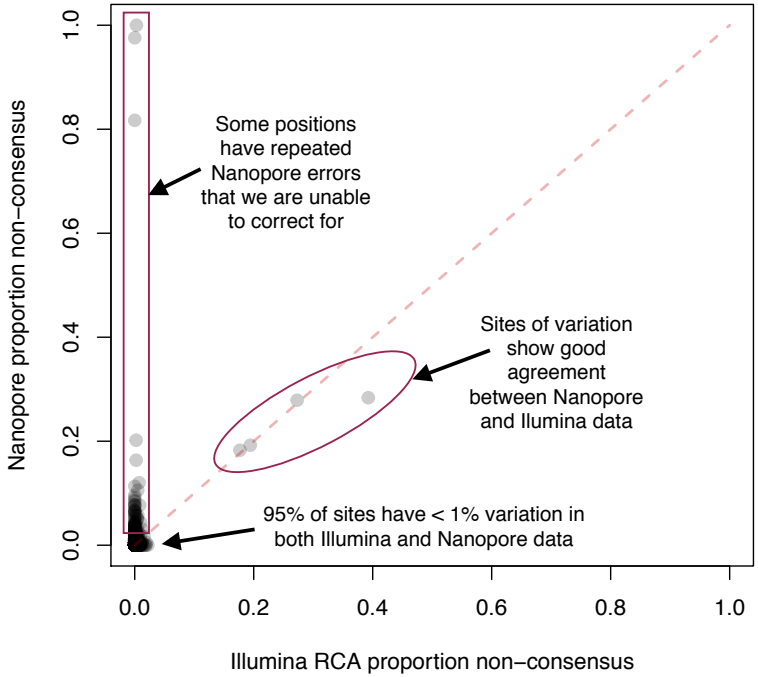
Philippa Matthews
Anna McNaughton
David Bonsall
Hannah Roberts
Mariateresa de Cesare
Paolo Piazza
Anthony Brown
Azim Ansari
Rory Bowden
Eleanor Barnes

HBV Workflows



Nanopore sequencing read error rates remain a challenge for calling within-sample variants.





Viromics as a diagnostic tool

Virus-agnostic Library + Comprehensive Enrichment Panel

- RNA-and-DNA libraries
- Curated list of viruses and bacteria (full-length, partial, rMLST)
- SureSelect RNA baits (~5Mb)
- For unknown samples
- For any included pathogen (e.g. EBV, HCMV, VSV)

Collaboration:

GAinS	**Both/Other**
Cyndi Goh	Azim Ansari
Julian Knight	Mariateresa de Cesare
Eduardo Svoren	Hubert Slawinski
Charles Hinds	David Bonsall
ChiMES	Amy Trebes
Tanya Golubchik	Paolo Piazza
Ivo Elliott	Anthony Brown
Andrew Pollard	Senthil Chinnakannan
Manish Sadarangani	Camilla Ip
Martin Maiden Group	Martin Maiden Group
Ellie Barnes	Ellie Barnes
Rory Bowden	Rory Bowden

Bacteria

Family	Species
Streptococcaceae	*Streptococcus pneumoniae*
	Streptococcus pyogenes
	Streptococcus agalactiae
Staphylococcaceae	*Staphylococcus aureus*
Mycoplasmataceae	*Mycoplasma pneumoniae*
Legionellaceae	*Legionella pneumophila*
Coxiellaceae	*Coxiella burnetii*
Enterobacteriaceae	*Escherichia coli*
	Klebsiella pneumoniae
	Klebsiella oxytoca
	Enterobacter cloacae
	Enterobacter aerogenes
	Serratia marcescens
Pasteurellaceae	*Haemophilus influenzae*
	Haemophilus parainfluenzae
Chlamydiaceae	*Chlamydophila pneumoniae*
	Chlamydia psittaci
Pseudomonadaceae	*Pseudomonas aeruginosa*
Moraxellaceae	*Moraxella catarrhalis*
Moraxellaceae	*Acinetobacter baumannii*
	Acinetobacter calcoaceticus
Mycobacteriaceae	*Mycobacterium tuberculosis*
Xanthomonadaceae	*Stenotrophomonas maltophilia*
Alcaligenaceae	*Bordetella pertussis*
Neisseriaceae	*Neisseria meningitidis*
Listeriaceae	*Listeria monocytogenes*
Spirochaetaceae	*Borrelia burgdorferi*
Spirochaetaceae	*Treponema pallidum*
Leptospiraceae	*Leptospira (multiple spp)*
Bartonellaceae	*Bartonella henselae*
Brucellaceae	*Brucella (multiple spp)*

Viruses

Adenoviridae
- Mastadenovirus A
- Mastadenovirus B
- Mastadenovirus C
- Mastadenovirus D
- Mastadenovirus E
- Mastadenovirus F
- Mastadenovirus G

Arenaviridae
- Lassa mammarenavirus
- Lymphocytic chriomeningitis mammarenavirus

Bunyaviridae
- California Encephalitis Virus
- Rift Valley Fever Virus
- Sandfly Fever Naples Virus
- Sandfly Fever Sicillian Virus

Coronaviridae
- Human Coronavirus HCoV-229E
- Human Coronavirus HCoV-NL63
- Human Coronavirus HCoV-HKU1 Genotypes A, B, C
- MERS-Coronavirus
- Human Coronavirus HCoV-OC43 Genotypes A-E
- SARS-Coronavirus

Flaviviridae
- Dengue Fever Virus Genotype 1
- Dengue Fever Virus Genotype 2
- Dengue Fever Virus Genotype 3
- Dengue Fever Virus Genotype 4
- Japanese Encephalitis Virus - All Genotypes
- Murray Valley Encephalitis Virus - All Genotypes
- St. Louis Encephalitis Virus - All Genotypes
- West Nile Virus - All Genotypes
- Tick-borne Encephalitis Virus - All Genotypes
- Yellow Fever Virus - All Genotypes

Herpesviridae
- HHV1 / Herpes Simplex Virus Type 1 (HSV-1)
- HHV2 / Herpes Simplex Virus Type 2 (HSV-2)
- HHV3 / Varicella-Zoster Virus (VZV)
- HHV4 / Epstein-Barr Virus (EBV)
- HHV5 / Human Cytomegalovirus (HCMV)
- HHV6A / Human Herpesvirus 6A
- HHV6B / Human Herpesvirus 6B
- HHV7 / Human Herpesvirus 7
- HHV8 / Kaposi's Sarcoma Herpesvirus (KSHV)

Orthomyxoviridae
- Influenza A virus - H1N1
- Influenza A virus - H1N2
- Influenza A virus - H2N2
- Influenza A virus - H3N2
- Influenza A virus - H5N1
- Influenza A virus - H7N3
- Influenza A virus - H7N7
- Influenza A virus - H7N9
- Influenza A virus - H9N2
- Influenza B virus
- Influenza C virus

Paramyxoviridae
- Human Parainfluenza virus 1
- Human Parainfluenza virus 3
- Human Parainfluenza virus 2
- Human Parainfluenza virus 4a
- Human Parainfluenza virus 4b
- Human Parainfluenza virus 5
- mumps virus - A
- mumps virus - B
- mumps virus - C
- mumps virus - D
- mumps virus - F
- mumps virus - G
- mumps virus - H
- mumps virus - I
- mumps virus - J
- mumps virus - K
- mumps virus - L
- mumps virus - N
- measles virus - A
- measles virus - B1
- measles virus - B2
- measles virus - B3
- measles virus - C1
- measles virus - C2
- measles virus - D1
- measles virus - D2
- measles virus - D3
- measles virus - D4
- measles virus - D5
- measles virus - D6
- measles virus - D7
- measles virus - D8
- measles virus - D9
- measles virus - D10
- measles virus - D11
- measles virus - E
- measles virus - F
- measles virus - G1
- measles virus - G2
- measles virus - G3
- measles virus - H1
- measles virus - H2
- sosuga virus
- hendra virus
- henipavirus - B
- henipavirus - M
- Respiratory syncytial virus - A
- Respiratory syncytial virus - B
- Human metapneumovirus

Parvoviridae
- Primate erythroparvovirus 1
- Primate tetraparvovirus 1
- Human bocavirus 1

Picornaviridae
- Human Parechovirus 1
- Human Parechovirus 2
- Human Parechovirus 3
- Human Parechovirus 4
- Human Parechovirus 5
- Human Parechovirus 6
- Human Parechovirus 7
- Human Parechovirus 8
- Parechovirus B
- Enterovirus B
- Enterovirus A
- Rhinovirus A
- Rhinovirus B
- Rhinovirus C
- Enterovirus D
- Cardiovirus A
- Cardiovirus B
- Cardiovirus B
- Cardiovirus B
- Cardiovirus B
- Hepatitis A
- Rosavirus 2
- Salivirus A
- Salivirus FHB

Polyomaviridae
- JC polyomavirus
- BK polyomavirus

Rotavirus
- Rotavirus A
- Rotavirus B
- Rotavirus C

Rhabdoviridae
- Rhabdovirus 1 - Rabies
- Rhabdovirus 4 - Duvenhage
- Rhabdovirus 5 - European Bat Lyssavirus 1 (EBLV2)
- Rhabdovirus 6 - European Bat Lyssavirus 2 (EBLV1)
- Rhabdovirus 7 - Austrailian Bat Lyssavirus(es)
- Rhabdovirus 2 - Lagos Bat virus
- Rhabdovirus 3 - Mokola virus

Togaviridae
- Eastern Equine Encephalitis Virus
- Western Equine Encephalitis Virus
- Venezuelan Equine Encephalitis Virus
- Rubella virus

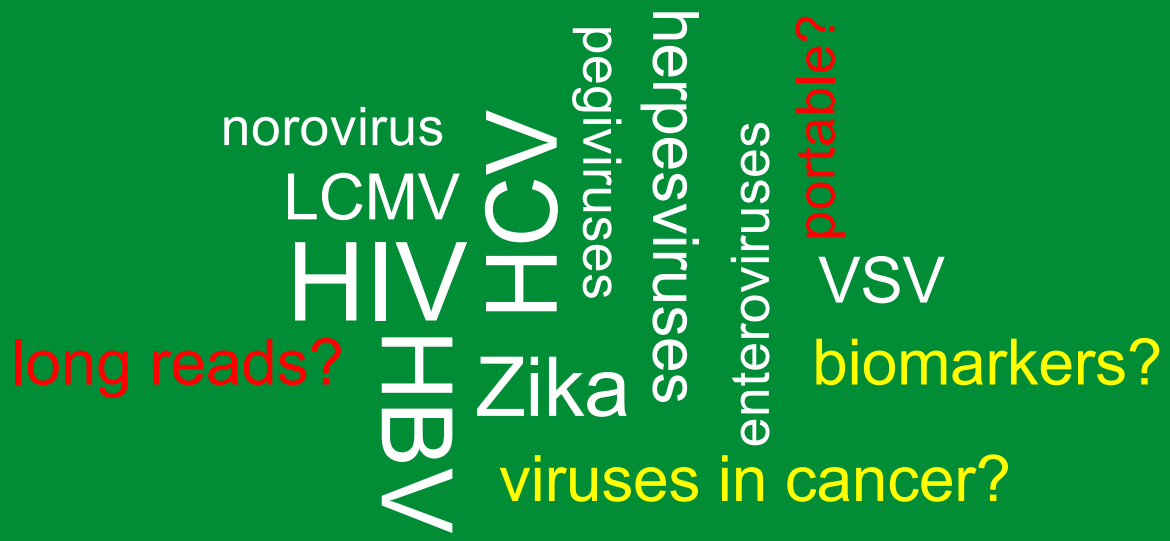
There are two types of metagenomics:

(1) Hay classification

https://hackingmaterials.com/2013/11/11/why-hack-materials

(2) Needle detection

U.S. Marine Corps photo by Lance Cpl. James Purschwitz



Contact us:
Rory Bowden rbowden@well.ox.ac.uk
David Bonsall david.bonsall@bdi.ox.ac.uk
Tanya Golubchik tanya.golubchik@bdi.ox.ac.uk
Mariateresa de Cesare decesare@well.ox.ac.uk
Azim Ansari azim.ansari@well.ox.ac.uk
Hannah Roberts hroberts@well.ox.ac.uk