

Massively Parallel Sequencing (MP-Seq™): a New Tool for Adventitious Agent Detection, Genetic Stability Evaluation and Clone Selection

David Onions PhD FMedSci DVMS(hon) FRSE
Chief Scientific Officer, BioReliance



 BioReliance
YOU CAN RELY ON US

Massively Parallel Sequencing

- Massively Parallel sequencing (Mp-Seq) is a new GMP method to identify all adventitious viruses, mycoplasma and bacteria and to characterise the state of the transcriptome. But MP-Seq is an adjunct not a replacement to traditional tests.
- In the presentation we focus on its ability to detect new viruses but in biotechnology applications it provides reassurance to regulators and sponsors that their product is safe!
- Data have been submitted to the FDA and EMA using this technology

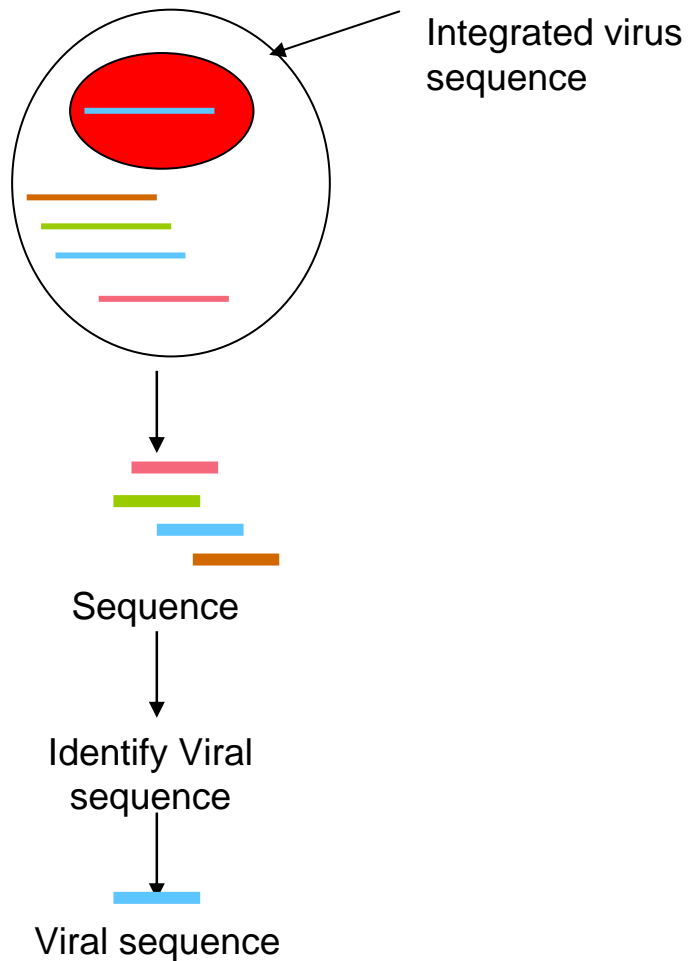
Massively Parallel Sequencing

- ~1 million reads each 300 to 400 bases per O/N run
- Virus detection two main approaches
 - Detection of virions
 - Detection of integrated & latent viruses
- BioReliance has developed proprietary algorithms and processes to support detection of viruses by massively parallel sequencing

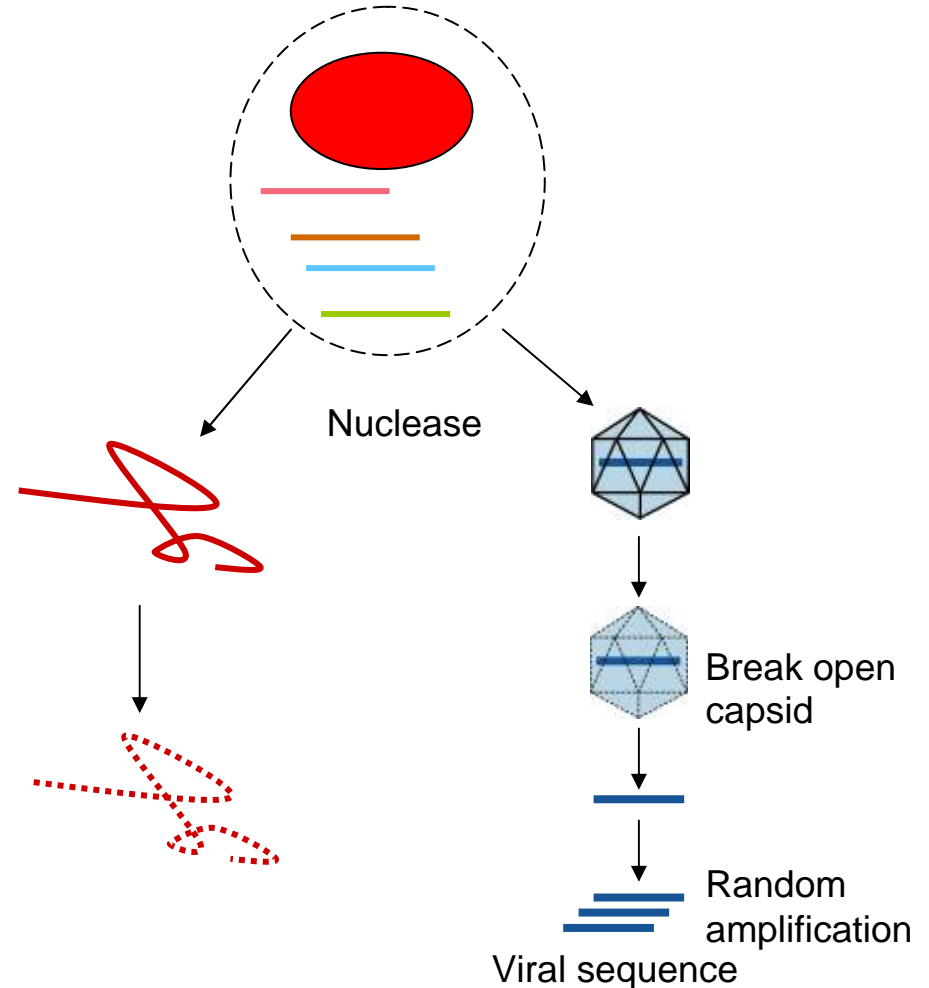


MP-Sequencing cell bank & supernatant off cells & serum

Deep Sequencing Transcriptome



Deep Sequencing Supernatant



Case study 2: bovine serum

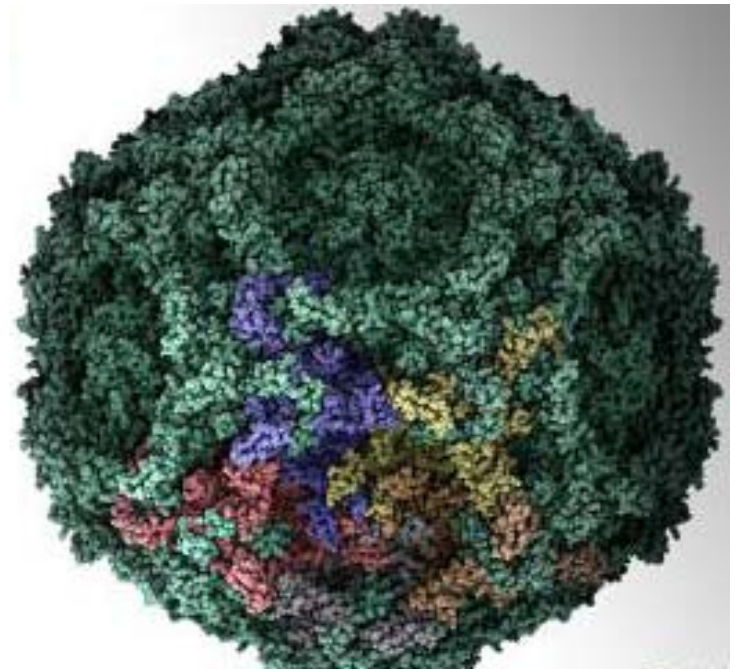
- Encapsidated BVDV and BPyV are prevalent in FBS but they are not the commonest viruses
- MP-sequencing of serum
 - BPV-3* >10,000 hits
 - BPV-2 >8,000 hits
 - **And 5 hits against****

*Allander, et al. . (2001) *Proc. Natl. Acad. Sci. USA* 98, 11609-11614

**Onions D Kolman J Massively parallel sequencing, a new method for detecting adventitious agents. *Biologicals*. 2010 377-80. 2010

New contaminant in bovine serum

- We have identified a new parvovirus in bovine serum (BAAV-2) using massively parallel sequencing
- It is able to infect human cells and cells of other species
- It can establish latent infections. Therefore cells that have been exposed to serum in the past need to be screened.
- It is a dependovirus (AAV) and is likely to be mobilised by adenoviruses and herpesviruses



New Bovine Parvoviruses

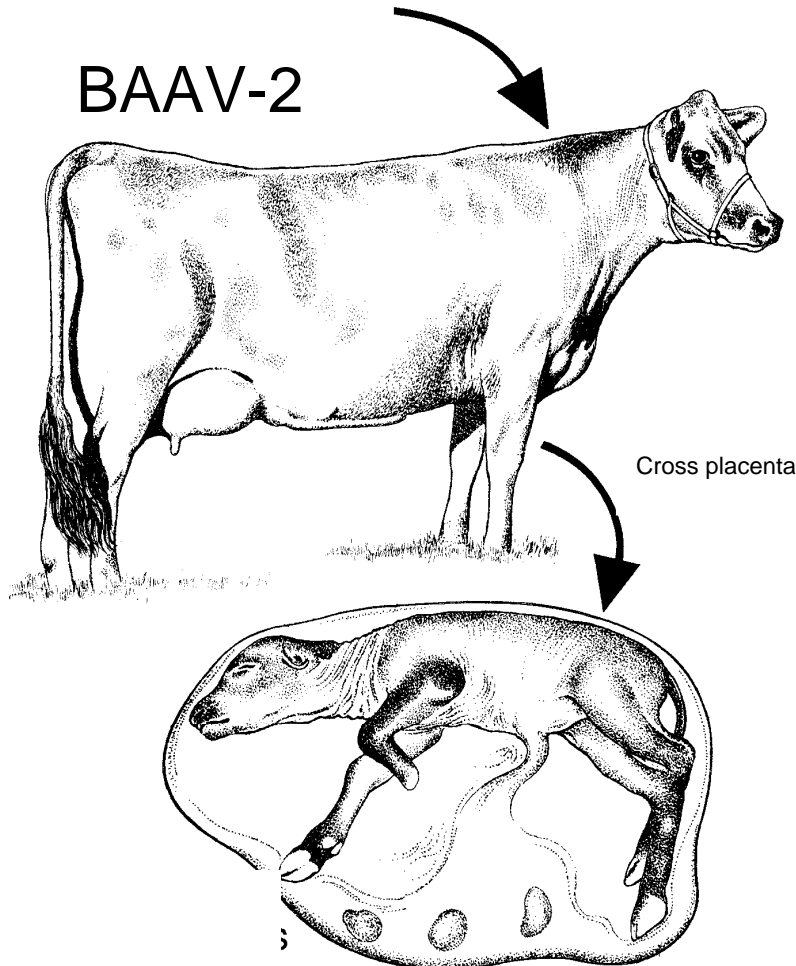
BPV-2

BPV-3

BAAV-2

BioReliance QPCR Data

(number positive number tested. Different sera than tested by MP-Seq)



Serum	BPV-2	BPV-3	BAAV-2
Calf	2/5	1/5	0/5*
FBS	2/5	3/5	1/5**

BPV-2 up to 6.7×10^3 ge/ml

BPV-3 up to 1.1×10^4 ge/ml

BAAV-2 < 100 ge/ml

Plus *2 positive below LOQ of 100 copies, **2 positive below LOQ

Other new viruses in serum

- Picornaviridae
 - When MP-seq of serum was analyzed at a lower e-value a potentially **new genotype of a known member of the *Picornaviridae* family, a *kobuvirus*, was identified.**
 - This virus has never been observed in serum before
 - But it was first detected as a contaminant of HeLa cells

This is an example of the less common contaminants like calicivirus (vesivirus) 2117 or Cache Valley virus. These are rare but potentially devastating contaminants

Case study 2 (cont'd): new contaminants in bovine serum

- In the last few years 4 new bovine parvoviruses have been detected in serum using MP-Seq methods
- At least one of these viruses has a wide host range and can establish latent infections
- These viruses are not covered by routine tests or regulatory guidance

Cell Banks exposed in their development to FBS should be screened for these viruses

But it doesn't affect me does it, I have an AOF process?

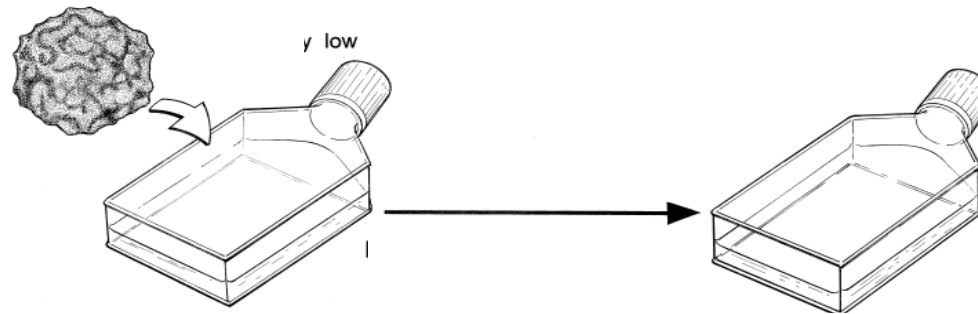
This is where plant peptones come from



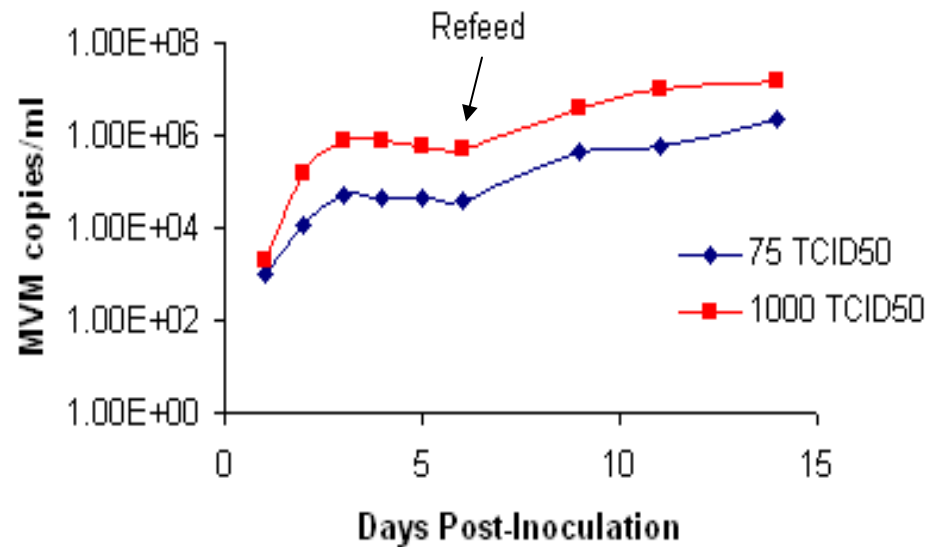
Parvoviruses are shed in faeces and are amongst the most resistant viruses in the environment!

New Applications: Rapid *in vitro* assay with MP-seq endpoint

CHO cells infected at low MOI, 1 in 200,000 cells

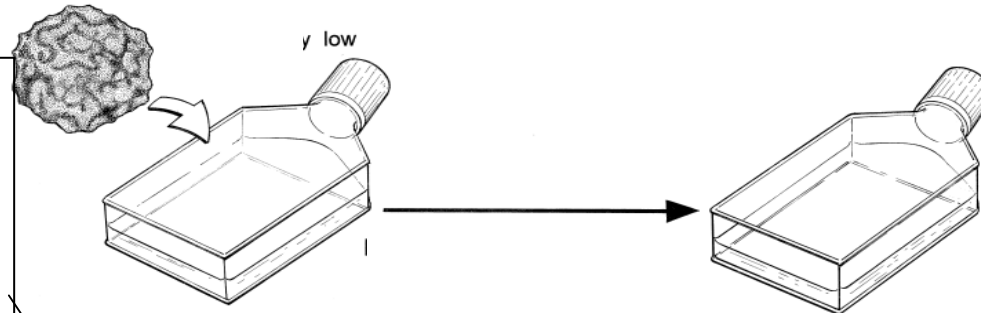


MVM Replication in CHO Cells

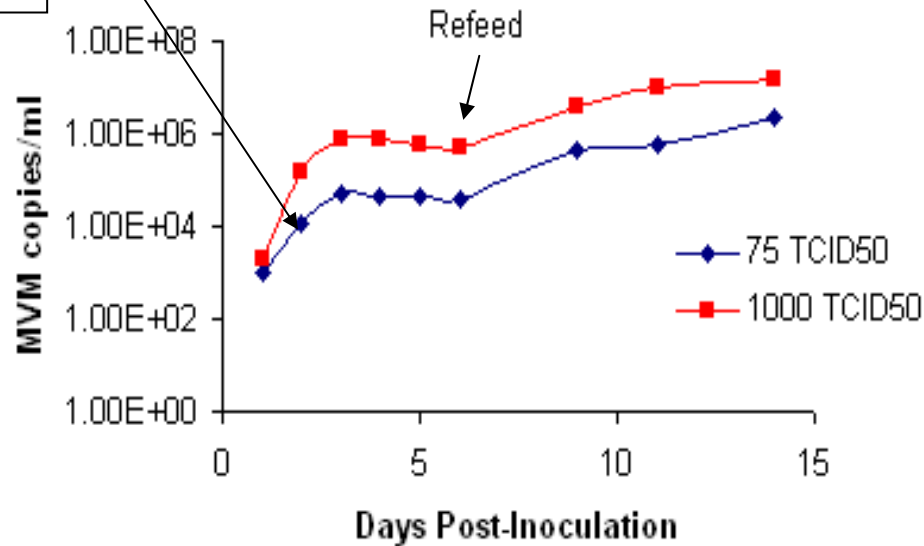


Rapid *in vitro* assay with MP-Seq endpoint

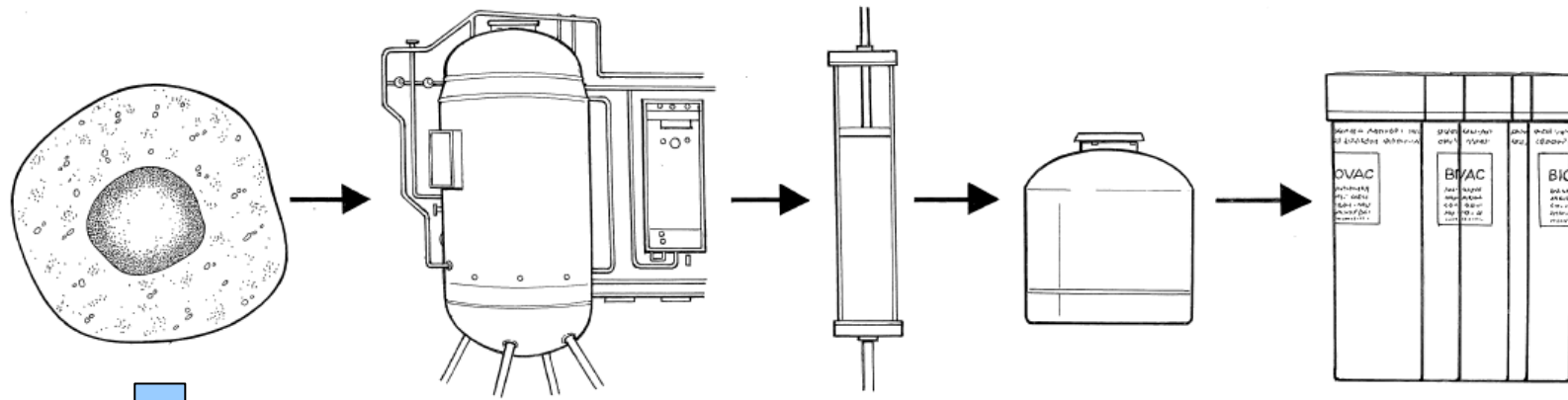
**MP-Seq
detects
Infection
day 4
before CPE**



MVM Replication in CHO Cells



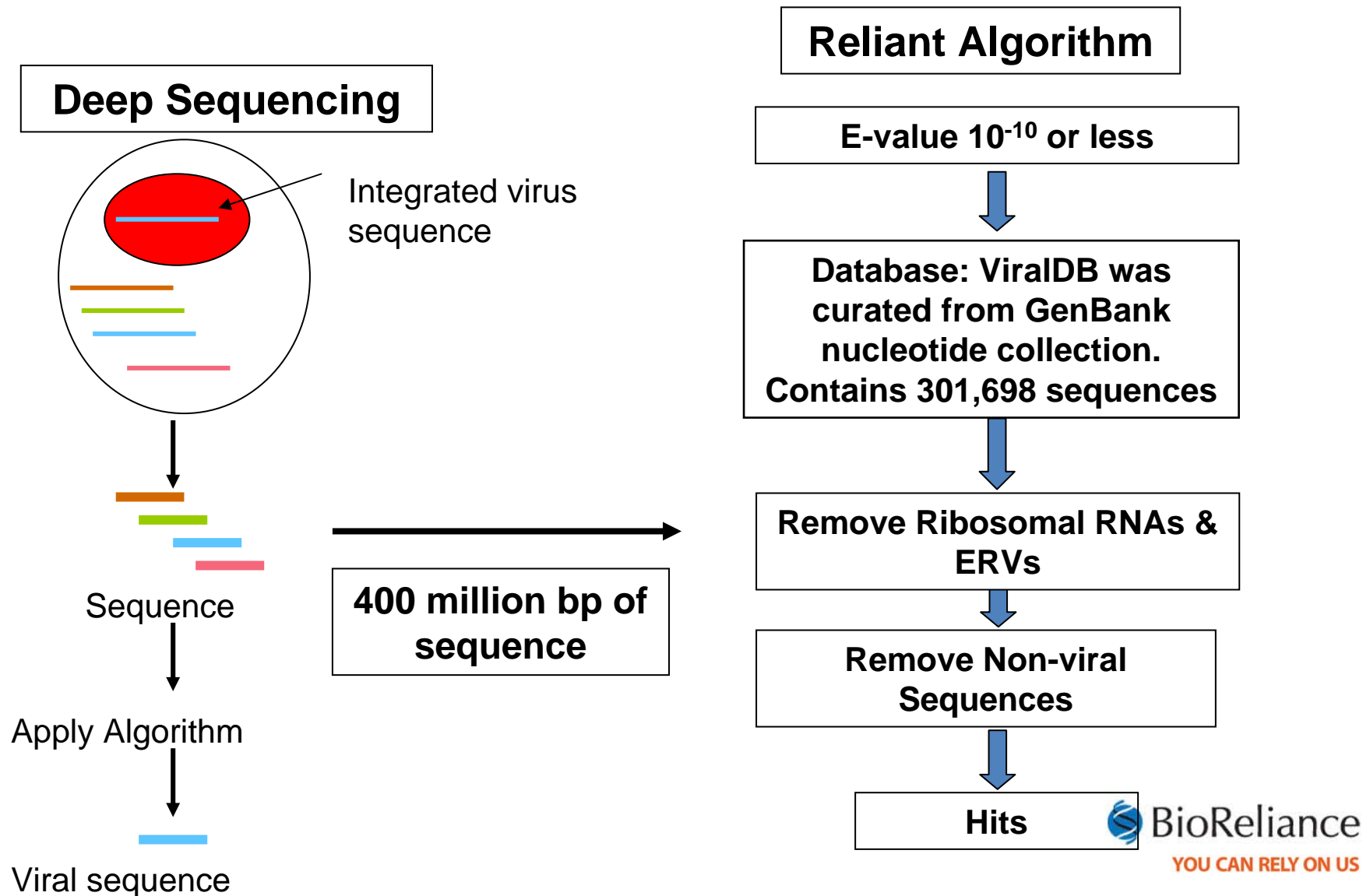
Transcriptome Analysis Case Studies



- ↓
- CHO
 - MRC 5
 - Vero
 - Hi 5
 - Sf9
 - Avian cells

**Transcriptome Analysis Is
Used to Verify The Safety
of Cell banks**

Case study 3: Screening MRC-5 cells



Why you need bioinformatics

400 Mbp output printed on A4 sheets would be 7.5 metres high!

7.5 metres

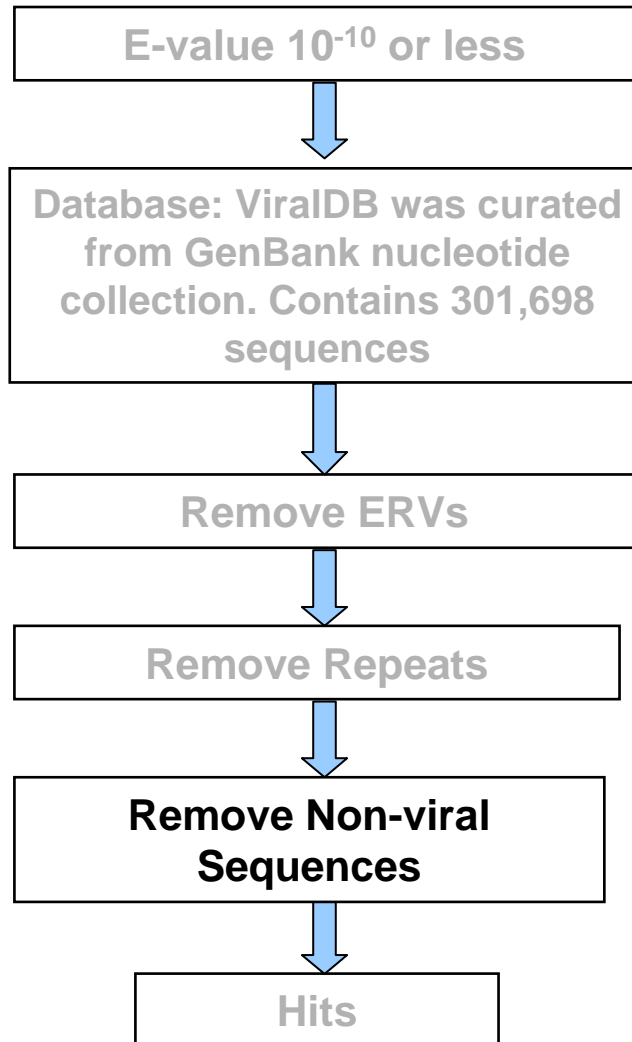
64,000 pages



Bioinformatics

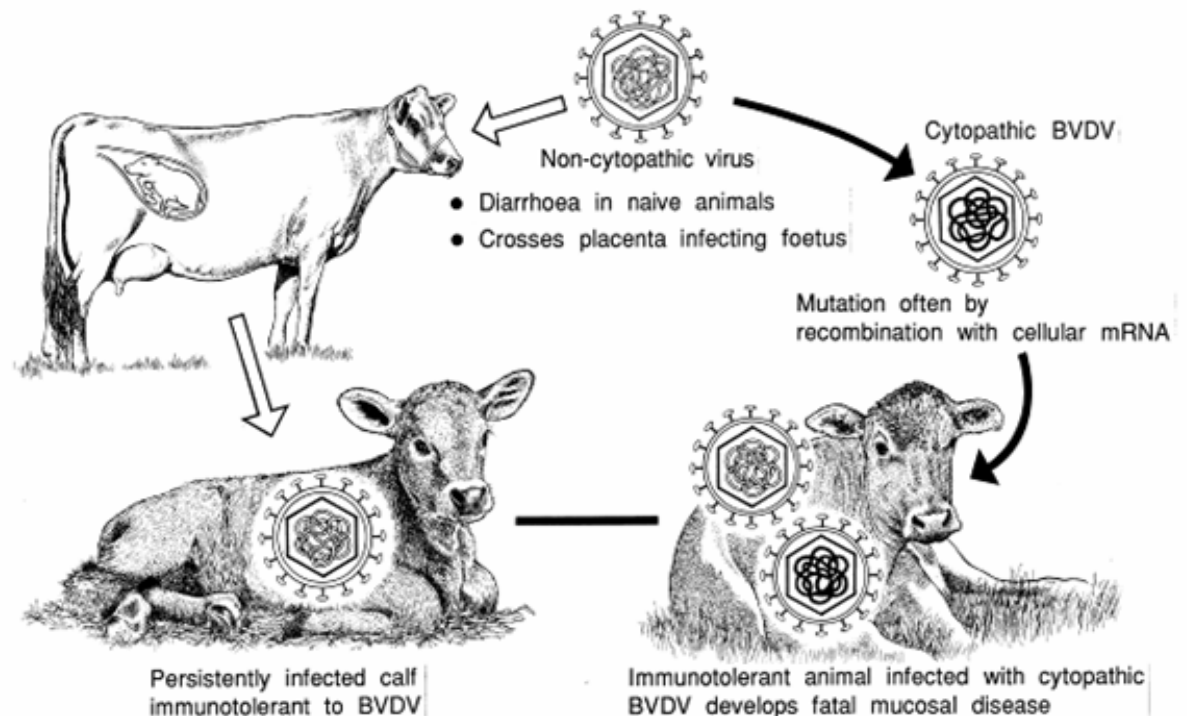
10 pages
Report

Algorithm to filter hits

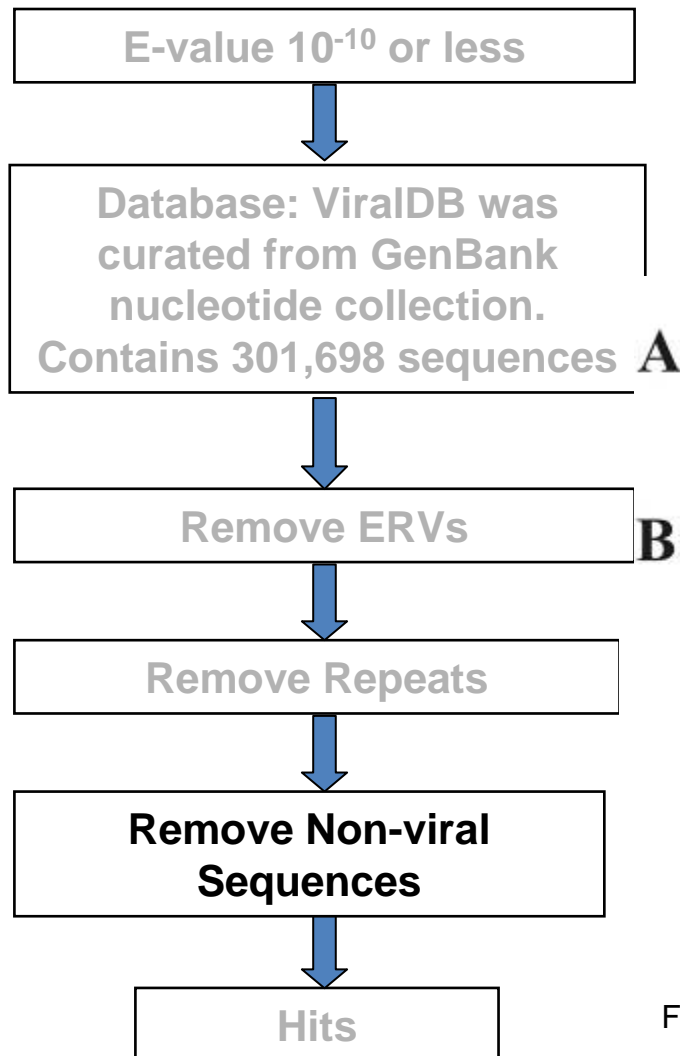


False hits: de novo recombination between virus and cellular sequences e.g. retroviruses and pestiviruses

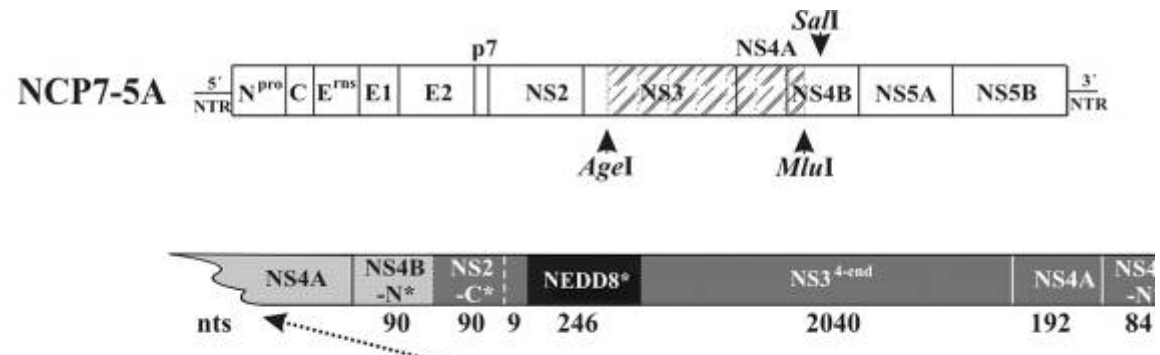
BOVINE VIRAL DIARRHOEA VIRUS



Algorithm to filter hits



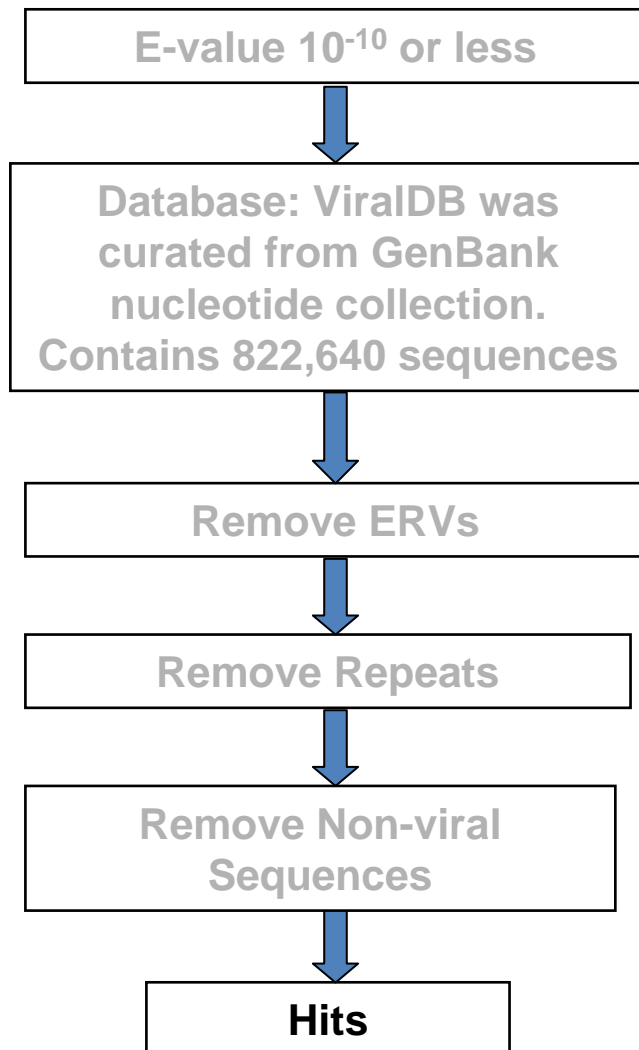
False hits: De novo recombination between virus and cellular sequences. e. g. retroviruses and pestiviruses



(A) Non-cytopathic BVDV

(B) Cytopathic BVDV has a large duplication of viral sequences together with an insertion encoding part of cellular Nedd8

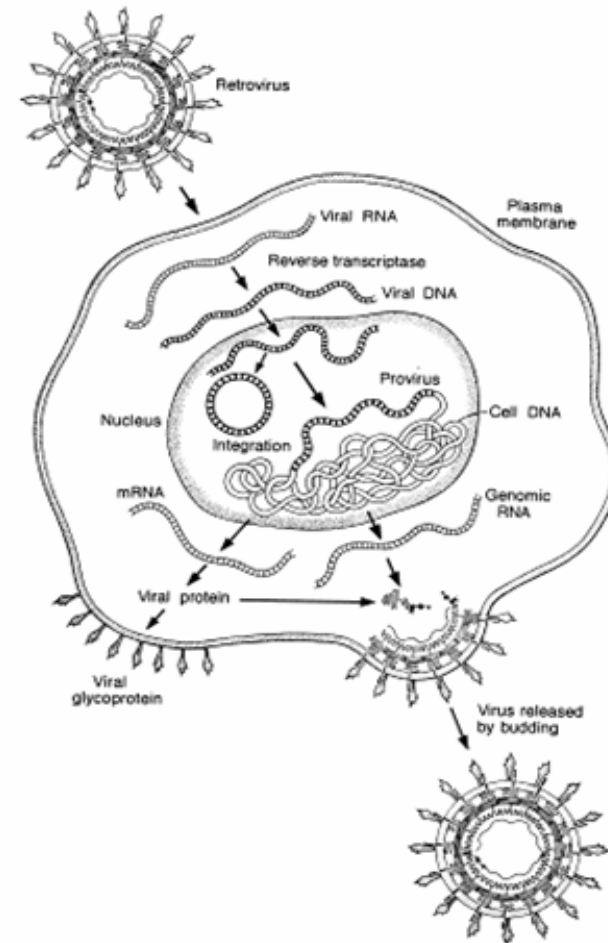
MP-Seq Case Study 3 (cont'd): MRC-5 transcriptome



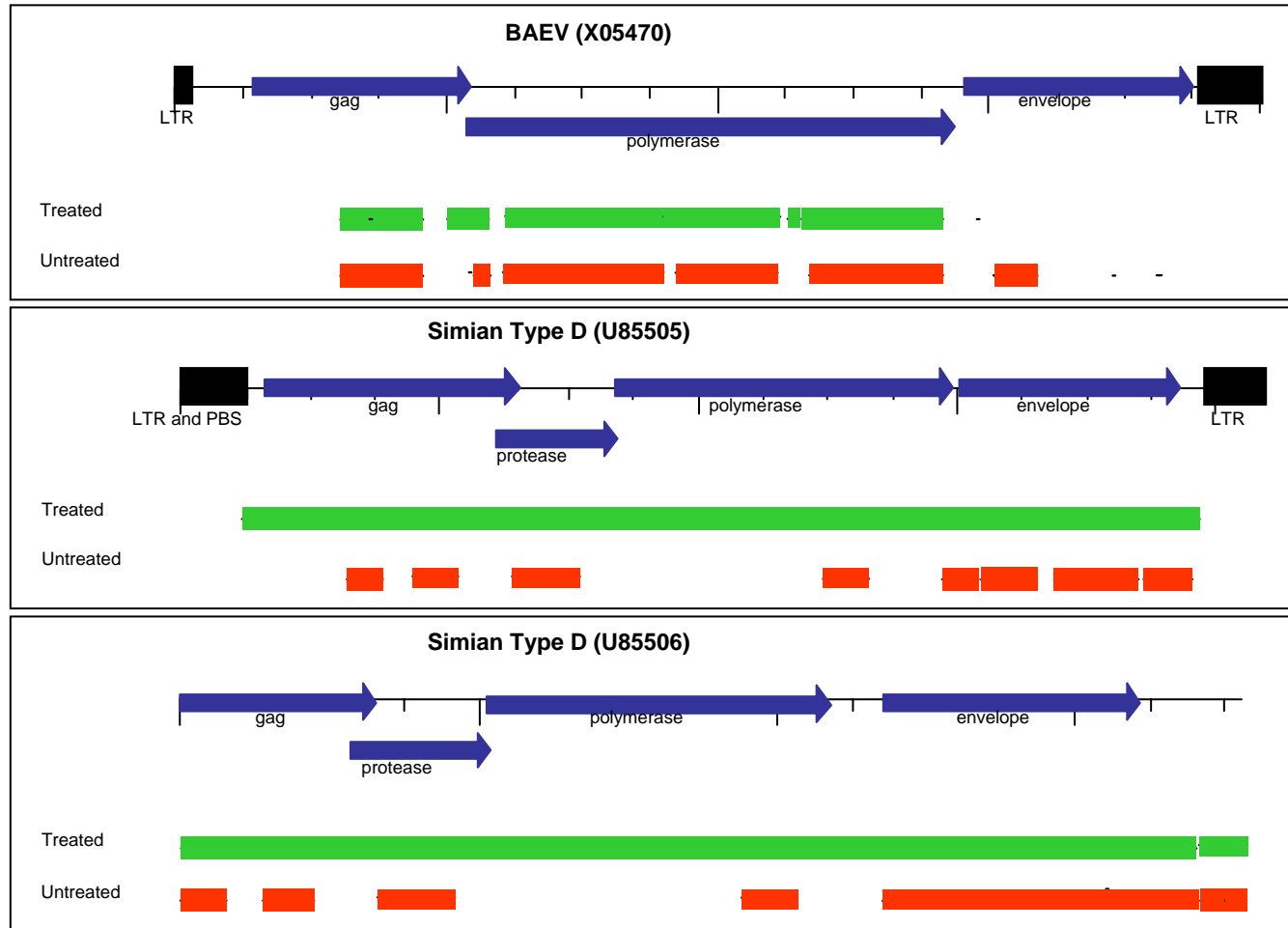
- 195 Million bp
- 746,844 reads
- 5,139 reads BLAST to ViralDB, of these
 - 1691 are HERVs
 - 2326 are false hits to BVDV (ubiquitin or related)
 - Remainder ribosomal RNAs
- Control breadth of coverage
 - B2M, GAPD, CGI-119, L37a, CALM2, S11, S13, OAZ1
- Control sensitivity
 - TUBB ~400c/cell HIT
 - CTBP1 ~300c/cell HIT
 - GOLGA1 ~100c/cell HIT

Case study 4: Vero Cell Analysis

- Vero cells analysed before and after treatment with inducing agents for retroviruses (IDU)
- Transcriptome analysis undertaken
- Porcine Circovirus was not detected (and confirmed by specific PCR)



Expressed Retrovirus Sequences in Vero cells



Dr Khan's group at the FDA has recently published similar data: Ma et al J. Virol 2011

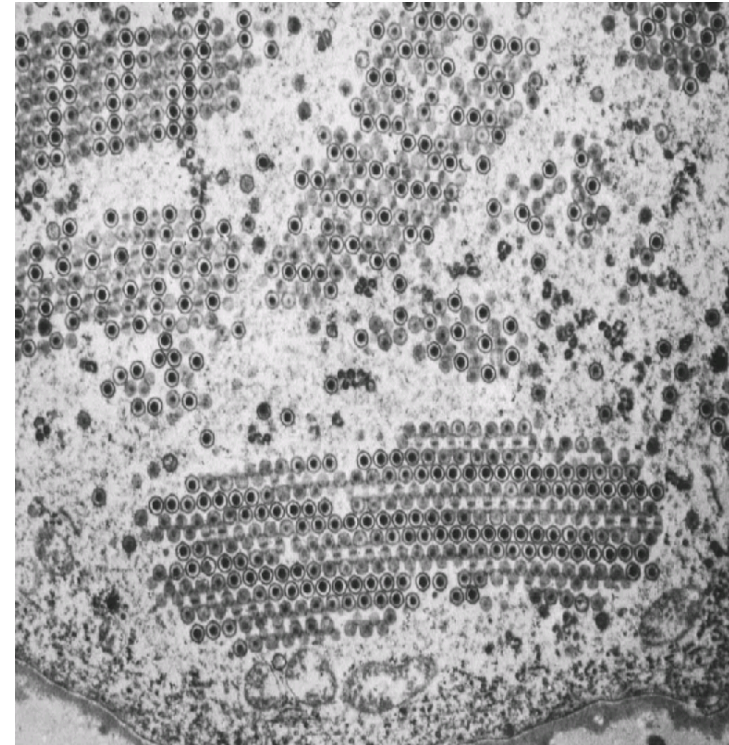
Case study 5: Insect cell transcriptome analysis

	Date Run	Treatment	Total reads	Ribosomal	Non-ribosomal	Unique Virus	% Hits Virus	P-Value
High 5	All Combined	Untreated	468,579	253,667	214,912	470	0.22%	0.4039
		Heat Shock	365,299	225,062	140,237	326	0.23%	
SF9	All Combined	Untreated	923,314	221,781	701,533	1,525	0.22%	1.6e ⁻³⁰
		Heat shock	863,111	218,400	644,711	2,060	0.32%	

Hit	# of Hits to Accession in DB		Seq Length in DB	Heat Shock			Untreated			Description
	Treated	Untreated		Coverage	Coverage (bp)	Coverage (%)	Coverage	Coverage (bp)	Coverage (%)	
AF254789	80	173	1,006	1-1,006	1,006	100.00%	1-1,006	1,006	100.00%	Autographa californica nucleopolyhedrovirus mutant vsk-1d1, genomic sequence [Autographa californica nucleopolyhedrovirus]

Case study 6: Nodavirus contamination of Hi 5 cells

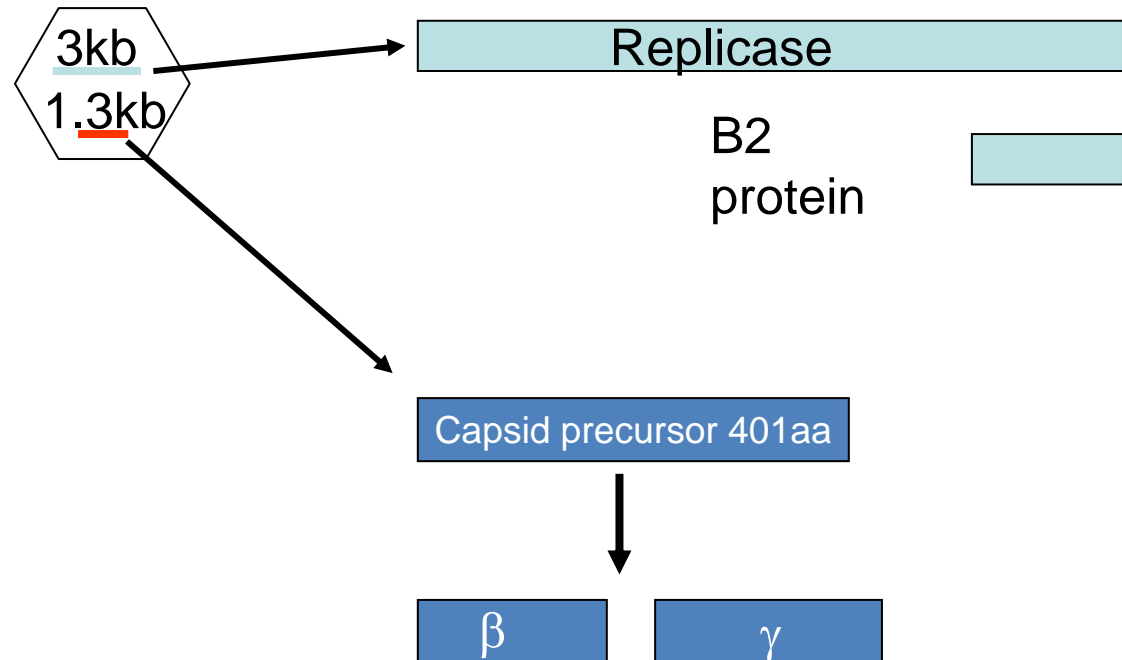
- Virus like particles found in Hi 5 cells
- No cytopathology on the Hi 5 cells
- Subsequently demonstrated to be a nodavirus by PCR
- Nodaviruses not detected in mammals but:
 - paracitio RNA virus can replicate in BHK cells and
 - Nodamura virus is lethal by ic inoculation in mice



Micrographs by George Reid
BioReliance

MPS easily detects the Nodavirus

- 1009 hits against nodaviruses
- Complete bipartite genome detected in Hi 5 cells
- Complete viral genome sequence assembled

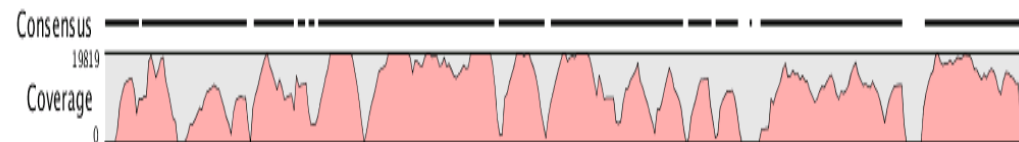
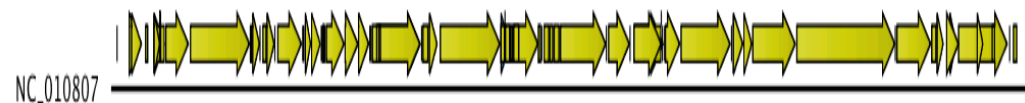


Case Study 7: bacteriophage contamination in *E. Coli*

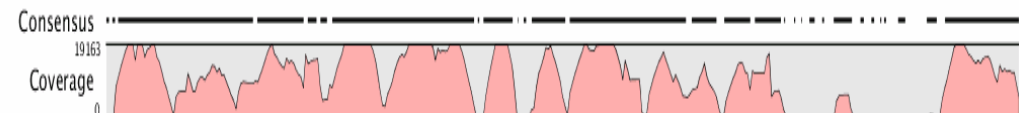
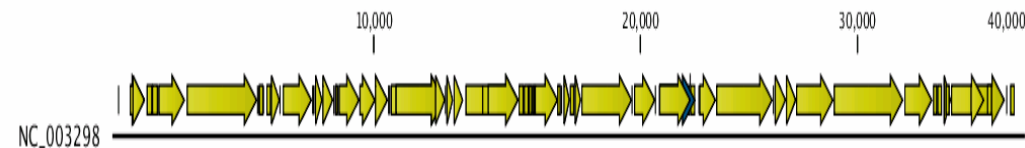
- Sponsor reported a catastrophic failure in a fermenter 4 hours after initiation
- A large proportion of *E. coli* noted to be non-viable
- BioReliance fermenter crash program initiated:
 - MP-Seq analysis of supnt.
 - Electron microscopy
 - Assays for lytic phage

MP-Seq generated >680,000 reads against bacteriophage T7 & T3 average length >400bp

EU547803; Salmonella phage phiSG-JL2, complete genome; NC_010807



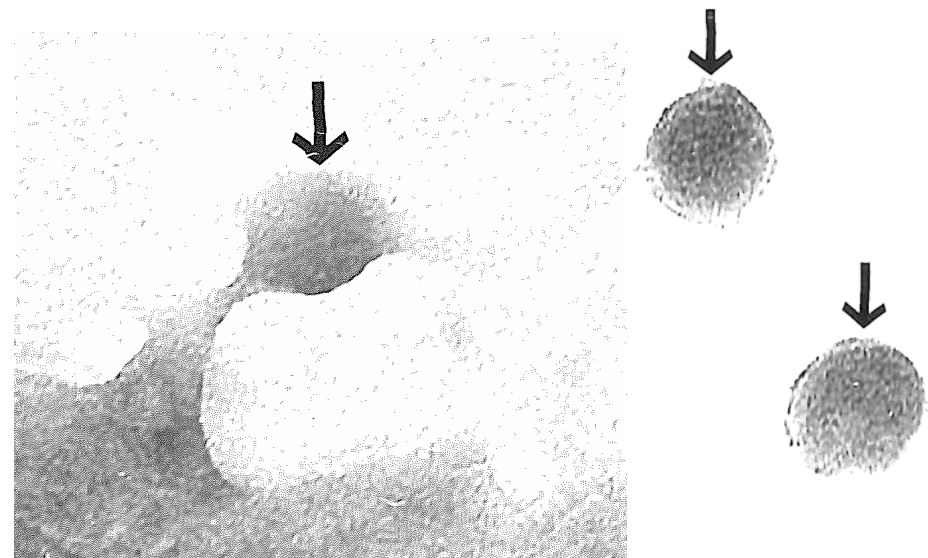
AJ18471; Bacteriophage T3, complete genome; NC_003298



Case Study 7: Bacteriophage contamination in E. Coli

- BioReliance fermenter crash program initiated:
 - MP-Seq analysis of supnt.
 - Electron microscopy
 - Assays for lytic phage
- Follow up actions
 - Specific PCRs to identify root cause of infection
 - Clean down with monitoring of phage sequences by PCR

Conclusion: A new member of the T7 phage family was the cause if the crash.
Phage presence confirmed by EM



Conclusion

- Massively parallel sequencing (MP-Seq) provides the start of the art method for verifying:
 - Freedom of cell substrates from latent or lytic adventitious agents
 - freedom of raw materials from contaminants
- Bioinformatics is the key component
- New applications:
 - Rapid endpoint in in vitro assays

Acknowledgements: John Kolman, Colette Cote, Brad Love & Audrey Chang