

Development of a GenomeTrakr Pipeline for Foodborne Viruses

Zhihui Yang, David Lacher, Mark Mammel

Division of Molecular Biology

Office of Applied Research and Safety Assessment,

CFSAN, FDA

09/27/2018

Foodborne illness acquired in the US – Major pathogens (CDC, 2011)

Pathogens	Foodborne illnesses/year	Note
Bacterial	3.6 million (39%)	
Parasitic & others	0.2 million (2%)	
Viral	5.5 million (59%)	Norovirus: 58%
Total (31 major pathogens*)	9.4 million	

- **There are an estimated 47.8 million foodborne illnesses in the US each year.**
- * **20% (9.4 million) were caused by 31 identified major pathogens (above).**
- **80% (38.4 million) of total foodborne illnesses were caused by unspecified agents:**
 - Known agents with insufficient data to estimate the agent-specific burden
 - Known agents not yet recognized as causing foodborne illness
 - Microbes, chemicals, or other substances known to be in food whose pathogenicity is unproven
 - Agents not yet described

Viruses with known and posited foodborne transmission

Clinical symptoms	Virus	Genome	Family	Routes of Transmission	Vaccine availability	References
Gastroenteritis	Norovirus	ssRNA(+) 7.3-8.3 kb	Caliciviridae	Fecal-oral	no	{Lopman, 2016}
	Rotavirus	dsRNA / 11 segments / 18,5 kb	Reoviridae	Fecal-oral	yes	{Grimwood, 2010}
	Adenovirus	dsDNA 35-36 kb	Adenoviridae	Fecal-oral / respiratory droplets / infected tissues and blood	yes	{Ison, 2016}
	Sapovirus	ssRNA(+) 7.1-7.7 kb	Caliciviridae	Fecal-oral	no	{Oka, 2015}
	Astrovirus	ssRNA(+) 6.8-7.0 kb	Astroviridae	Fecal-oral	no	{Bosch, 2014}
	Aichi virus	ssRNA(+) 8.3 kb	Picornaviridae	Fecal-oral	no	{Kitajima, 2015}
Hepatitis	HAV	ssRNA(+) 7.5 kb	Picornaviridae	Fecal-oral / blood	yes	{Vaughan, 2014}
	HEV	ssRNA(+) 7.2 kb	hepeviridae	water / zoonosis / fomite ?	no	{Cook, 2015} {Nieuwenhuijse, 2017}
Neurological diseases	Poliovirus	ssRNA(+) 7.5 kb	Picornaviridae	Fecal-oral	yes	{Mueller, 2005}
	Nipah virus	ssRNA(-) 18 kb	paramyxoviridae	food / zoonosis / animal bite	yes	{Luby, 2006}
Others	HPAI virus H5N1	ssRNA(-) / 8 segments / 13.5 kb		water / zoonosis	yes	{Tumpey, 2002}
	SARS coronavirus	ssRNA(+) 27-32 kb	Coronaviridae	Fecal-oral / respiratory/zoonosis	no	{Newell, 2010}{Lau, 2005}
	Coxsackievirus	ssRNA(+) 7.2 - 8.5 kb	Picornaviridae	Fecal-oral	no	{Ma, 2011}
	(Hemorrhagic fever, respiratory disease, myocarditis...)	other non-polio Enterovirus(rhinovirus, echovirus, numbered enteroviruses)	ssRNA(+) 7.2 - 8.5 kb	Picornaviridae	Fecal-oral / respiratory zoonosis / body fluids / food?	no
	Ebolavirus ?	ssRNA(-) 18-19 kb	Filoviridae		under development	{Mann, 2015}
	Torque teno virus	circular ssDNA(-) 3.8 kb	Anelloviridae	blood / fecal-oral	no	{Spandole, 2015} {Pinho-Nascimento, 2011}

Yang, Z. *IJFM* 2017.

(More than 40% of the 130+ known viruses that can infect humans → foodborne transmission)

FDA NCBI BioProject for Foodborne Viruses

NCBI Resources ▾ How To ▾ Sign in to NCBI

BioProject Help

[Create alert](#) [Advanced](#) [Browse by Project attributes](#)

Display Settings: ▾

Send to: ▾

foodborne viruses

Accession: PRJNA396739 ID: 396739

Genome sequencing of foodborne viruses by FDA

Foodborne viruses are the leading cause of foodborne illnesses outpacing all other pathogens combined including bacteria. This project, initiated by the Molecular Virology Team in the Division of Molecular Biology/ Office of Applied Research and Safety Assessment, aims to (1) sequence a wide range of foodborne viruses (both commonly documented and less commonly documented) from clinical specimens, food and water; and (2) thus provide additional reference sequences for phylogenetic analysis and epidemiologic studies linked to foodborne illnesses.

Accession	PRJNA396739
Type	Umbrella project
Submission	Registration date: 1-Aug-2017 OARSA/CFSAN/FDA
Relevance	Foodborne viruses

Related information

- BioProject
- Data projects

Recent activity

- [Turn Off](#) [Clear](#)
- foodborne viruses BioProject
 - PRJNA396739 (1) BioProject
 - Gyrase A Mutations in Campylobacter Associated with Decreased Susceptibi PubMed
 - Use of whole-genome sequencing for Campylobacter surveillance from NAR PubMed
 - Whitehouse CA (96) PubMed

foodborne viruses encompasses the following 3 sub-projects:

Project Type		Number of Projects
Umbrella project		3
BioProject accession	Name	Title
PRJNA433976	Metagenomic sequences from food	Metagenomic sequences from food (OARSA/CFSAN/FDA)
PRJNA433975	Metagenomic sequences from human clinical samples	Metagenomic sequences from human clinical samples (OARSA/CFSAN/FDA)
PRJNA433977	Metagenomic sequences from water	Metagenomic sequences from water (OARSA/CFSAN/FDA)

IMPACT:

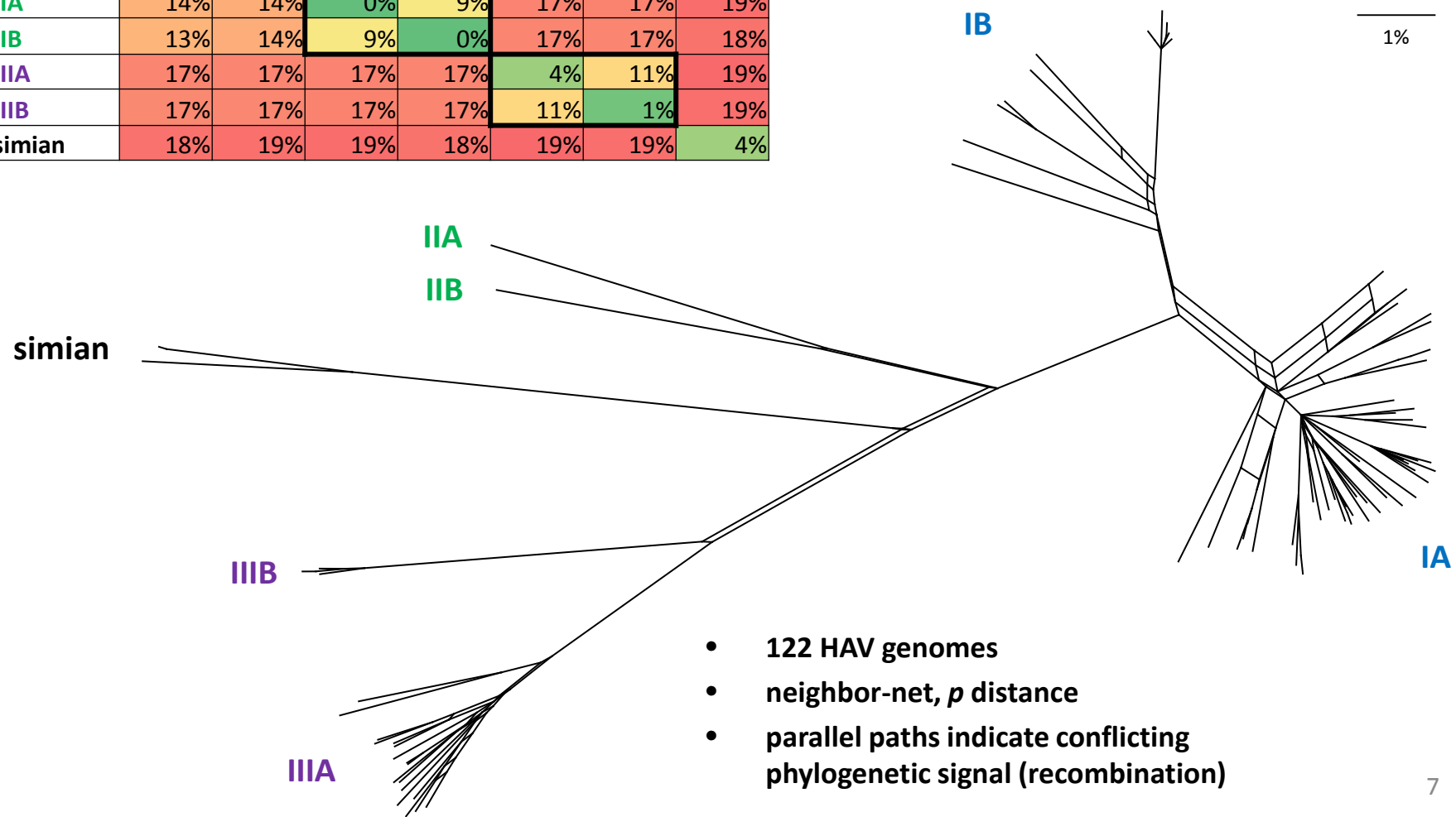
- Sequence database of a wide range of foodborne viruses**
- Publicly accessible**
- Phylogenetic analysis**
- Epidemiologic studies**
- Monitor potential outbreaks**

Identification of reads from WGS by kmer database

sample: E14S1_120_dpi_S2	reads	%
Picornaviridae_Hepatovirus_Hepatovirus A	624372	19.57
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IA	236	0.02
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IA_cluster1	19188	0.49
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IA_cluster2	108	0.01
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IA_cluster3	21	0.01
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IB	71444	13.98
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IB_cluster1	515806	65.89
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IB_cluster2	81	0.01
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIA	2	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIB	9	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIIA	14	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIIA_cluster1	6	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIIA_cluster2	7	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Hepatitis A virus IIIB	17	0.00
Picornaviridae_Hepatovirus_Hepatovirus A_Simian hepatitis A virus	1	0.00

Hepatitis A

Genotype	IA	IB	IIA	IIB	IIIA	IIIB	simian
IA	4%	9%	14%	13%	17%	17%	18%
IB	9%	3%	14%	14%	17%	17%	19%
IIA	14%	14%	0%	9%	17%	17%	19%
IIB	13%	14%	9%	0%	17%	17%	18%
IIIA	17%	17%	17%	17%	4%	11%	19%
IIIB	17%	17%	17%	17%	11%	1%	19%
simian	18%	19%	19%	18%	19%	19%	4%



- 122 HAV genomes
- neighbor-net, *p* distance
- parallel paths indicate conflicting phylogenetic signal (recombination)

Thank you!