



Bioinformatics Analysis of the Epitope Regions for Noroviruses Capsid Protein

Liping Chen

Huzhou Center For Disease Control and Prevention

2012.10.13 Shanghai

Noroviruses

Frozen strawberries likely the source of Germany's food poisoning outbreak

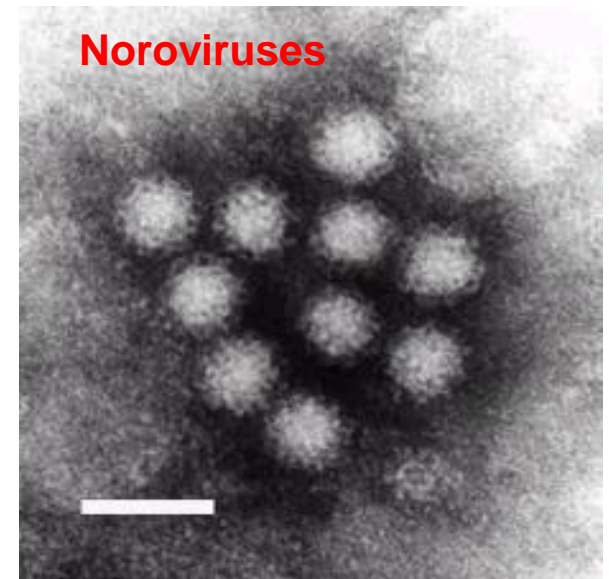
GERMANY | OCTOBER 5, 2012 | BY: ROBERT HERRIMAN | [+ Subscribe](#)



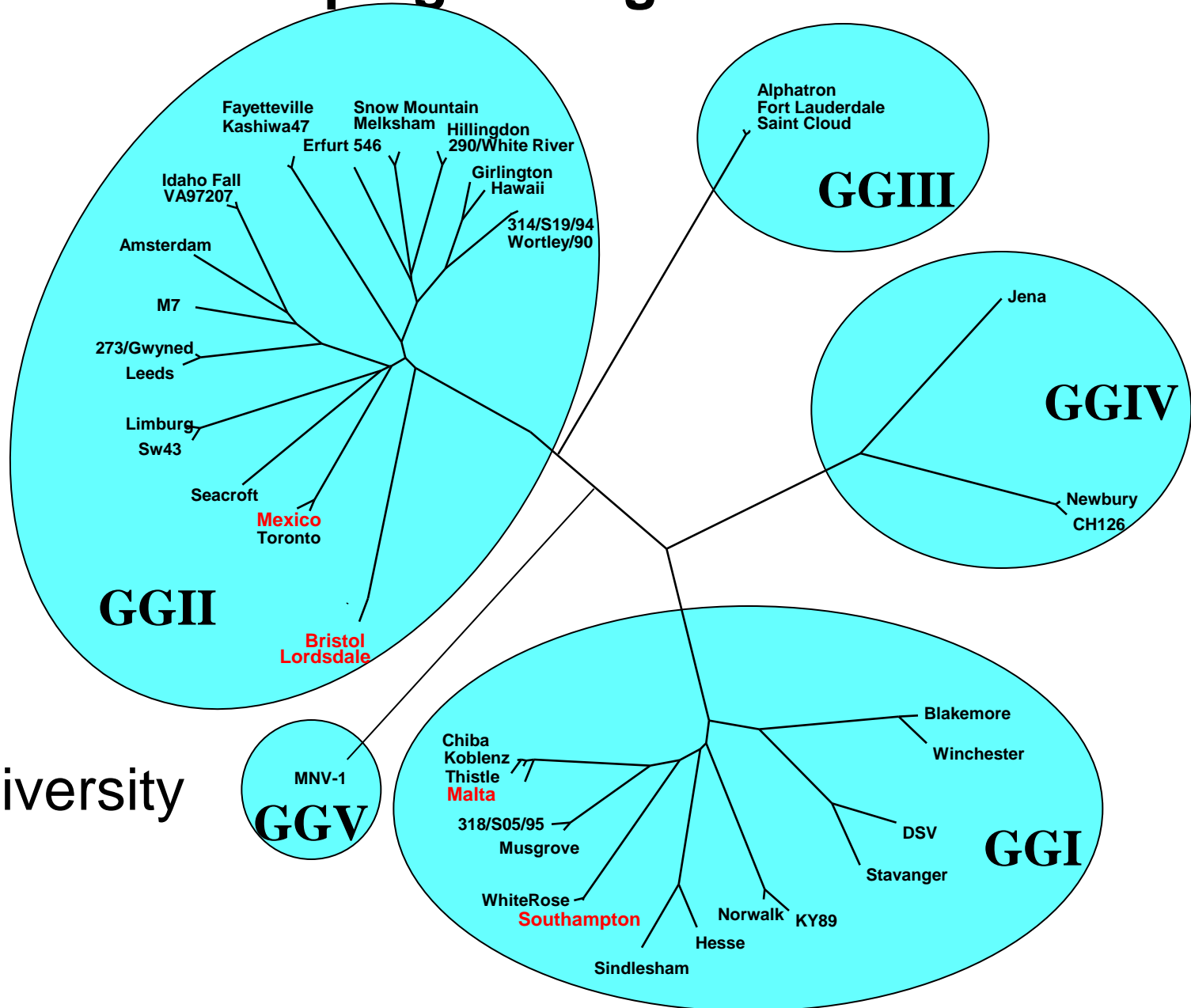
- Noroviruses are a group of viruses that cause the “stomach flu,” or gastroenteritis in people
- The symptoms include nausea, vomiting, diarrhea, and some stomach cramping.

Noroviruses

- **Family : *Caliciviridae***
- **Non-enveloped small round structured viruses (27-32 nm diameter)**
- **Predominantly epidemic**
- **The most common cause of outbreaks of gastroenteritis**

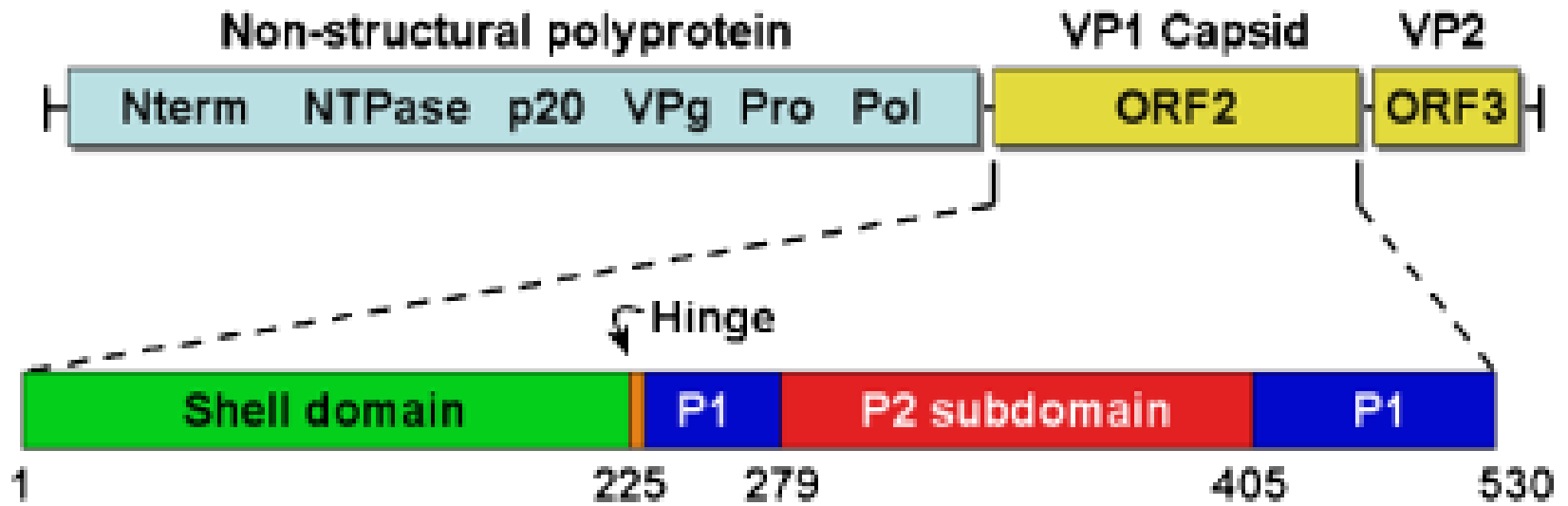
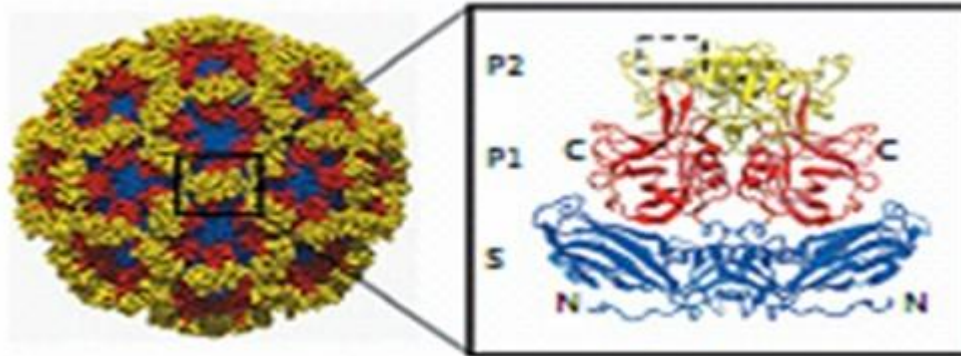


Phylogenetic Grouping among Noroviruses



- Virus diversity

The Genome of Norovirus



Current Diagnosis Methods

- PCR assays and real-time PCR assays
- electron microscopy
- enzyme Immunoassays

Fast diagnosis methods are needed!



Unique antibody?



Unique epitope?

Strategy

Sequences download and analysis



Protein 3D structure modeling



Conformational epitope prediction

Methods

- 18 full-length human norovirus capsid protein sequences were downloaded from NCBI
- Modeller9.9 was used for protein modeling
<http://www.salilab.org/modeller/9.9/release.html>
- SEPPA for predicting the possible spatial epitope
<http://lifecenter.sgst.cn/seppa/>

18 sequences

GI.1	ABW74128	Sweden	2004
	ACN32270	Sweden	2007
	AAS86780	USA	2001
GI.2	ACU56258	Belgium	2003
GI.3	ACX33982	Sweden	2007
GI.4	ACV41096	USA	2008
GI.8	ADB54834	USA	2008

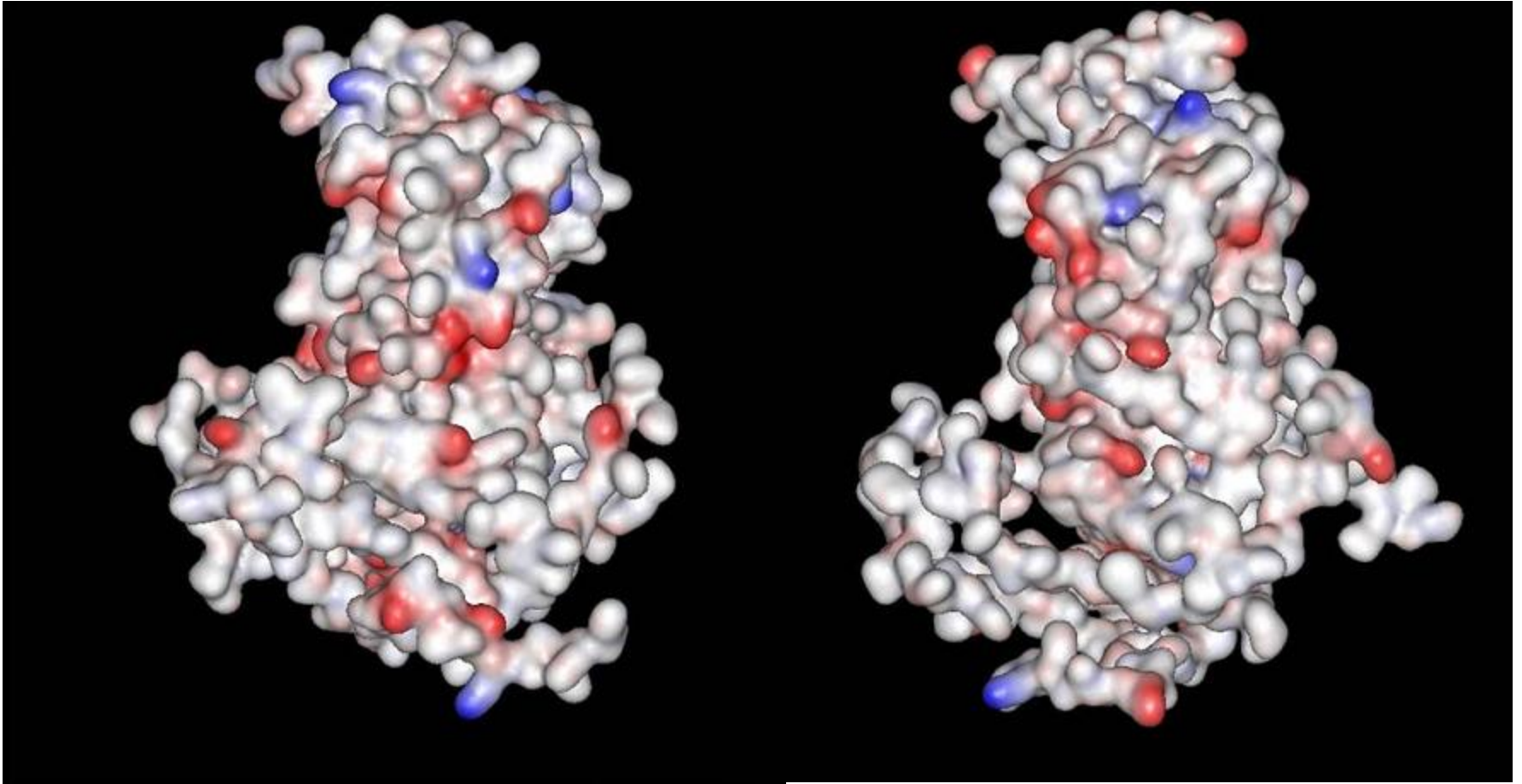
GII.1	AAL13016	USA	1994
GII.2	BAG68716	Netherlands	2005
GII.3	ADK23787	KOR	2006
GII.4	ABC96332	China	2001
	ABL74391	Netherlands	2006
	ABL74397	Netherlands	2006
	AEG79292	China	2010
GII.6	ADE28721	JPN	2008
GII.7	ACX85810	AUS	2008
GII.12	ADZ24003	USA	2010
GII.13	ACX81355	JPN	2008

Result of Sequences Analysis



- Four sites deletion of GII (14-16\28\192-193\530-531)
- unique insertion of GII3 and GII6 (304-319)
- unique insert fragment of GII4 at 417-421

Result of 3D Structures of Norovirus



GI

GII

Antigenic Prediction for AAL1301601.pdb:



Chain: A

Threshold: 1.80

Number of total residues: 535

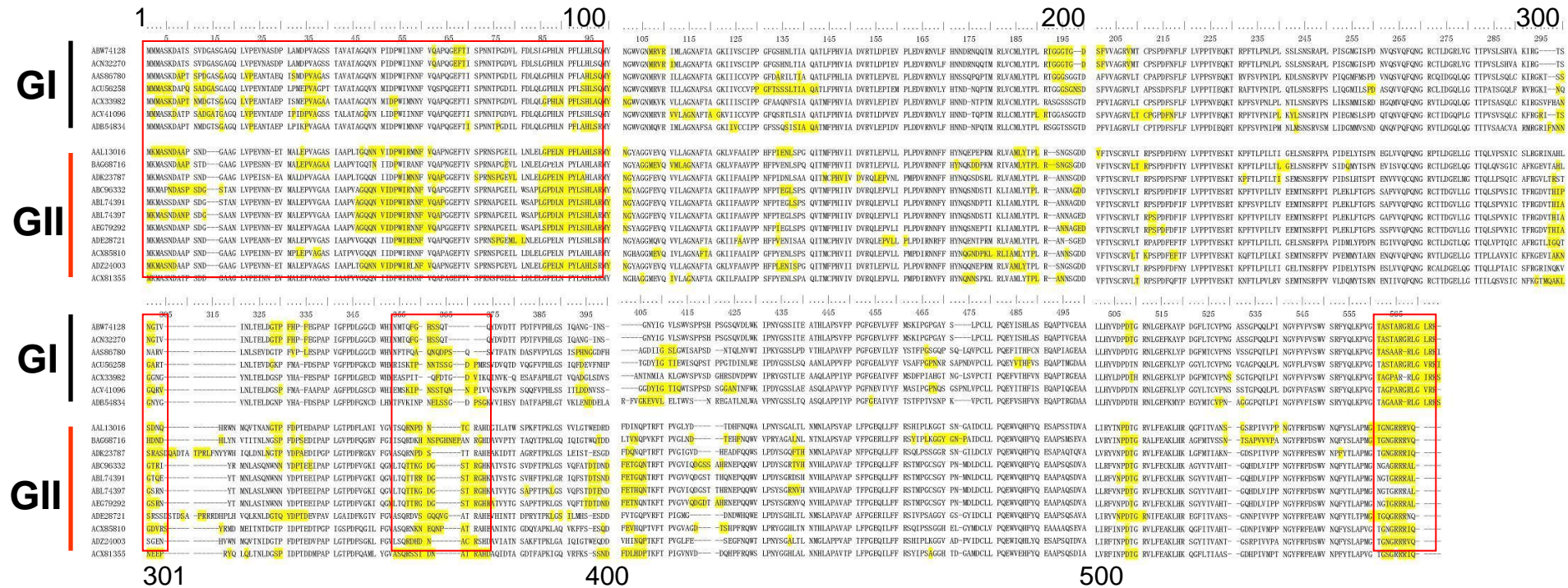
Number of predicted epitope residues: 77

[View 3D structure in Jmol](#)

1-50	M K M A S N D A A P S N D G A A G L V P E V N N E T M A L E P V A G A S I A A P L T G Q N N V I D P
51-100	W I R M N F V Q A P N G E F T V S P R N S P G E I L L N L E L G P E L N P F L A H L S R M Y N G Y A
101-150	G G V E V Q V L L a g n a F T A G K L V F A A I P P H F P I E N L S P G Q I T M F P H V i d V R T
151-200	L E P V L L P L p d v R n n f F h y n q E p E P R m R L V A M L Y T P L R s N G S G D D V F T V S C
201-250	r v l t R p S P D F d F N Y L V P P t v e S K t k p f T l p I L T I G E L S N s r f p a P i D E L Y
251-300	t S P N E G L V V Q p q n g r p T L D G E L l g t t Q L V P S N i c S L R G R i N A H L S D N Q H R
301-350	W N m Q V T N A N G T P F D P T E D a P a P L g t p d f l a N i Y g V t S Q R N P D N T C R A H D g
351-400	I l a t W s P K F T p K L G S v V L G T W E d R D F D I N Q P T R f T P V g L Y D T D H F N q W A L
401-450	p N Y S G A L T L N M N l a P S V A P L F P G E Q L l f f R s H I P L K g g t s n g a i d c l l p q
451-500	e W V Q H F Y q e s a P S S T D V a L i R y T N P D T G R V L F e a K L H R Q g f i t v A N S G S R
501-550	P I V V P P N g Y f R F D S W V N Q F Y S L A P M G T G N G R R R V Q

Predicted result format: **EPITOPE RESIDUE** | NON-EPITOPE RESIDUE | core residue

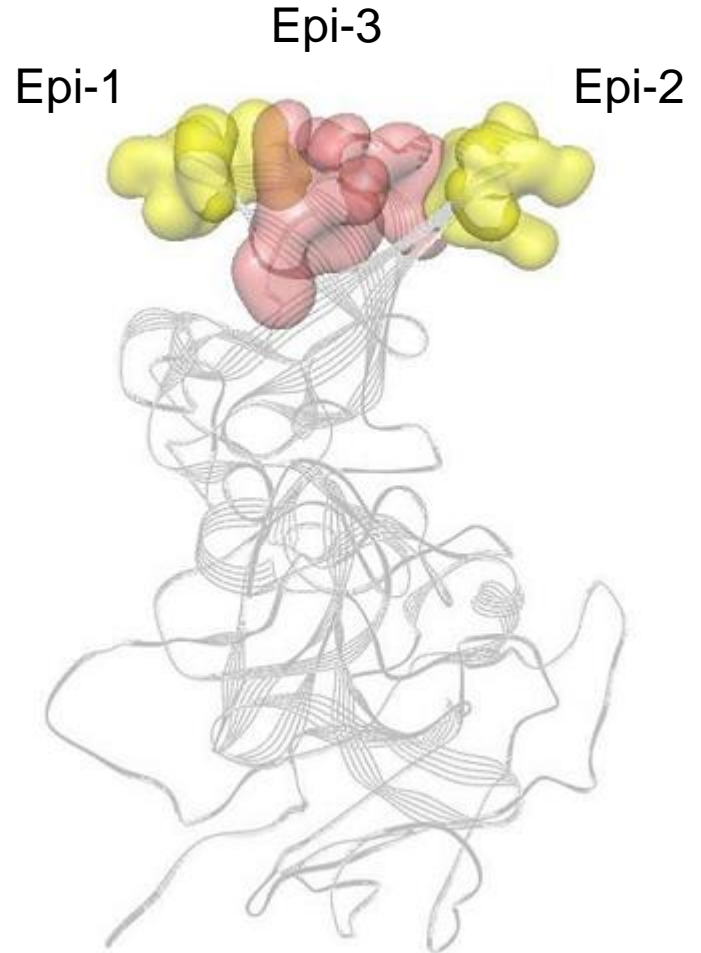
The Result of Predicted Epitope Regions



•The potential spatial epitope are highlighted in yellow

The Spatial Epitope of Norovirus Capsid Protein

305	315	325	335	345	355	365	375	385	395	405
NGIV	INLTELGGTP	FHP-PEGPAP	IGFPPDLGGCD	WHINMTQFG	RSSQT	QVDVDTI	PDTFYPHILGS	IQANG-INS-		GNVI
NGIV	INLTELGGTP	FHP-PEGPAP	IGFPPDLGGCD	WHINMTQFG	RSSQT	QVDVDTI	PDTFYPHILGS	IQANG-INS-		GNVI
NARY	LNLSEVDGTP	FYP-LBSAPAP	VGFPPDLGGCD	WHVNPYQQA	QNGDSS-Q	SVTPATN	DASFYPHILGS	ISPHNGGDFH		AGDII
GART	LNLTEVDGTP	FMA-FDSAPAP	VGFPPDPGKCD	WHMRISKTI	NTSSG-D	PMRSVDVQTD	VGFYPHILGS	IQFDEYFNP		IGDVI
GGNG	YNLTELGGSP	YHA-FDSAPAP	IGFPPDLGGCD	WHREASPIIT	QFDTG-D	VIKQINVK-Q	ESAFAPHLGI	VQADGLSDVS		ANTNM
GQKV	LNLTELGGSP	FMA-FAAPAP	AGFPPDLGGCD	WHLEMSKTI	SSSTQN-N	PIVVSVKPN	SQQFVPHLGS	ITLLEQVSS-		GGDEI
GNVG	YNLTELGGNP	YHA-FDSAPAP	LGFPPDPGKCD	LHMTFVKINP	NELSSG-D	PSGRVVIHSY	DATFAPHLGI	VKLENDDELA		R-FYQKEY
SDNQ	HIRMN	MQVINAQNTF	FDPTEDAPAP	LGTDFELANI	VGTVSQRNPD-N	TC	RAHIDGILATW	SPKFTPKLGS	VYLGITWEDRD	FDINQPTRF
HDND	HLYN	VITLNLGSP	FDPSIEDIPAP	LGVPDPGRKV	PGLISQBDAAH	NSPGHNEPAN	RGHDAVYPIY	TAQYTPRLGQ	IQIGTWEDD	LTYNQPVKF
SRSASQADTA	TPKLFNYWH	IQLDNLGTP	YDPAEDIPGP	LGTDPGRKV	PGVASQRNPD-S	TT	RAHEAKIDT	AGRFYPKLGS	LEIST-ESGD	FDQNGPTRF
GTRI	YR	MNLASQWNN	YDPTTELPAF	LGTDPVVKI	GGMLTQTTRG-DG	SI	RGRKATVSTG	SVDFYPKLGS	VQFAYTDND	FETGGNTRF
GTQE	YT	MNLASQWNN	YDPTTELPAF	LGTDPVVKI	GGVLTQTTRR-DG	SI	RGRKATVSTG	SVDFYPKLGR	IQFSTDTSD	FETGGNTRF
GSRN	YT	MNLASLWNN	YDPTTELPAF	LGTDPVVKI	GGVLTQTTRG-DG	SI	RGRKATVYTG	SAPFTPKLGS	VQFSTDTSD	FETHGNTRF
GSRN	YT	MNLASLWNN	YDPTTELPAF	LGTDPVVKI	GGMLTQTTRG-DG	SI	RGRKATVYTG	SAPFTPKLGS	VQFTTDND	FETHGNTRF
ESSSSTSDSA	PRRRDHPLH	VQLKDLGDTQ	YDPTDEVPAP	LGAIDFKGV	PGVASQRDVS	GGVVG	AT	RAHEVHINT	DPRTYPKLGS	ILMES-ESDD
GDVKS	YRMD	MEITNDGTP	IDPTEDTPGP	IGSPDPGGIL	PGVASQRNKN	HQNP	AT	RAHEANINTG	GDQYAPRLQA	VKFFS-ESDD
SGEN	HVRN	MQVINAQNTF	FDPTEDVPAP	LGTDPSPGKL	PGVLSQRDND-N	AC	RSHDAVIAIN	SARFTPKLGA	IQIGTWEDD	VHINQPTRF
NEEF	RYQ	LQLNLGSP	IDPTDMPAP	LGTDPGQML	VGVASQRSDI-DN	AT	RAHDAQIDTA	GDTFAPKIQG	VRFKS-SSND	FDLHNPTRF



Epi-1 Epi-2 Epi-3

Common epitopes for GI and GII: Epi-1\Epi-2
 Exclusive epitope for GII: Epi-3

Conclusion

- We find two common epitope regions for GI and GII genogroup, and an exclusive epitope region for GIII genogroup
- With the fast development of computational immunology tools, the bioinformatics will be more and more critical to vaccine design

Thanks!