

Quantifying Weather and Climate Impacts on Health in Developing Countries (QWeCI)



A Seventh Framework Programme Collaborative Project (SICA)

13 partners from 9 countries

www.liv.ac.uk/QWeCI

Grant agreement 243964

The validity of using automated evidence of pathogen occurrence from bibliometric sources

Marie McIntyre

Institute of Infection & Global Health,
University of Liverpool, Liverpool, U.K.

k.m.mcintyre@liv.ac.uk

UNILIV (8), UP(4), KNUST(5), CSE(2), IPD(2), UOC(2), ILRI(4),
UNIMA (9)



WP1.1 Disease database

- **Objective**

- Expanding the database (in conjunction with ENHanCE project) by the addition of vector species... and known climate and/or environmental sensitivities ...in an African context

- **Task**

- Task 1.1b : Use of existing UNILIV pathogen database with input from African based partners to add host, pathogen and vector data for diseases important in Africa

- **Deliverable**

- D1.1.a Report on current climate controls on selected infectious disease in Africa, based on database analysis and projection

- **Milestone**

- M1.1.a Completion of main phase of data entry to database

ENHanCED Infectious Diseases (EID2) database

Category

- [frequently pathogenic](#)
- [microbial rosetta stone database](#)
- [taylor organisms](#)
- [qweCI target organisms \(pathogens/vectors\)](#)
- [qweCI cargos with taxid](#)

Carriers Recursive rift valley fever virus

Name	Recursive	Our Rank	Our Division	NCBI Division	NCBI Rank
anopheles coustani	No	species	eukaryota	invertebrates	spec
bos taurus	No	species	eukaryota	mammals	spe
caeca hirus	No	species	eukaryota	mammals	spe
culicx pipiens	No	species	eukaryota	invertebrates	spe
homo sapiens	No	species	eukaryota	primates	spe
homo sapiens	Yes	species	eukaryota	primates	spe
mansonia africana	No	species	eukaryota	invertebrates	spe
ochlerotatus taeniorhynchus	No	species	eukaryota	invertebrates	spe
oris arnas	No	species	eukaryota	mammals	spe
rattus rattus	No	species	eukaryota	rodents	spe

Scientific Name	rift valley fever virus
Alternative Names	rvfv
Parent	phlebovirus
Path	click here
NCBI TaxId	11588
Class	viruses
NCBI Rank	species
NCBI Division	viruses
Categories	5
Recursive Carriers	9
Recursive Cargos	0
Recursive Climate Data Tab Separated	click here
Annual Average Temperature Histogram	click here
Annual Average Rainfall Histogram	click here
Recursive Countries	18
Recursive Countries KML	click here (please install Google Earth)
Recursive Countries JPEG Map	click here
Current Climate JPEG Map	Sorry no map yet.
Direct Papers	681
Direct Papers Time Series	click here
Latest Pubmed Papers	click here

Path

```

root
viruses
ssrna viruses
ssrna negative-strand viruses
bunyaviridae
phlebovirus
rift valley fever virus
    
```

Rift Valley fever virus clone 13 segment S, complete sequence

GenBank: DQ380182.1

[FASTA](#) [Graphics](#)

[Go to:](#) GenBank

LOCUS DQ380182 1141 bp RNA linear VRL 31-JAN-2007

DEFINITION Rift Valley fever virus clone 13 segment S, complete sequence.

ACCESSION DQ380182

VERSION DQ380182.1 GI:87622754

KEYWORDS .

SOURCE Rift Valley fever virus

ORGANISM [Rift Valley fever virus](#)

Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.

REFERENCE 1 (bases 1 to 1141)

AUTHORS Bird,B.H., Khristova,M.L. and Nichol,S.T.

TITLE Rift Valley fever virus genomics

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1141)

AUTHORS Bird,B.H., Khristova,M.L. and Nichol,S.T.

TITLE Direct Submission

JOURNAL Submitted (27-JAN-2006) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333, USA

FEATURES

source Location/Qualifiers

1..1141

/organism="Rift Valley fever virus"

/mol_type="genomic RNA"

/isolation_source="plaque pick of RVFV 74H559"

/db_xref="taxon:11588"

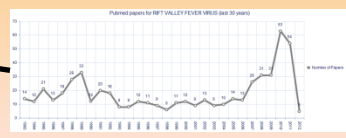
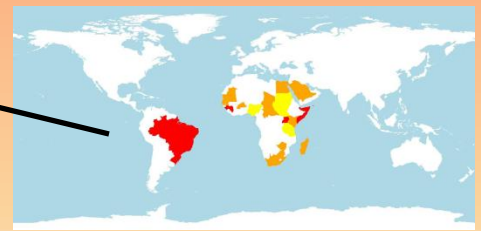
/segment="S"

/clone="13"

/country="Central African Republic"

/collection_date="1974"

primer_bind 1..22



<http://www.ncbi.nlm.nih.gov/nuccore/87622041>

GI identifier

Rift Valley fever virus strain 2250/74 segment S, complete sequence

GenBank: DQ380143.1

[FASTA](#) [Graphics](#)

[Go to:](#)

LOCUS DQ380143 1690 bp RNA linear VRL 07-NOV-2007
 DEFINITION Rift Valley fever virus strain 2250/74 segment S, complete sequence.
 ACCESSION DQ380143
 VERSION DQ380143.1 GI:87622041
 KEYWORDS .
 SOURCE Rift Valley fever virus
 ORGANISM [Rift Valley fever virus](#)
 Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
 REFERENCE 1 (bases 1 to 1690)
 AUTHORS Bird,B.H., Khristova,M.L., Rollin,P.E., Ksiazek,T.G. and Nichol,S.T.
 TITLE Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry
 JOURNAL J. Virol. 81 (6), 2805-2816 (2007)
 PUBMED [17192303](#)
 REFERENCE 2 (bases 1 to 1690)
 AUTHORS Bird,B.H., Khristova,M.L. and Nichol,S.T.
 TITLE Direct Submission
 JOURNAL Submitted (27-JAN-2006) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333, USA
 FEATURES Location/Qualifiers
 source 1..1690
 /organism="Rift Valley fever virus"
 /mol_type="genomic RNA"
 /strain="2250/74"
 /isolation_source="fetus"
 /host="bovine"
 /db_xref="GenBank:11588"
 /segment="S"
 /country="Zimbabwe: Beatrice"
 /collection_date="1974"
 primer_bind 1..22
 gene 35..832

Carriers Recursive [rift valley fever virus](#)

Name	Recursive	Our Rank	Our Division	NCBI Division	NCBI Rank	TaxId	Interaction Details
anopheles coustani	No	species	eukaryota	invertebrates	species	139045	click here
bos taurus	No	species	eukaryota	mammals	species	9913	click here
capra hircus	No	species	eukaryota	mammals	species	9925	click here
culex pipiens	No	species	eukaryota	invertebrates	species	7175	click here
homo sapiens	No	species	eukaryota	primates	species	9606	click here
homo sapiens	Yes	species	eukaryota	primates	species	9606	click here
mansonina africana	No	species	eukaryota	invertebrates	species	667564	click here
ochlerotatus taeniorhynchus	No	species	eukaryota	invertebrates	species	329105	click here
ovis aries	No	species	eukaryota	mammals	species	9940	click here
rattus rattus	No	species	eukaryota	rodents	species	10117	click here

<http://www.ncbi.nlm.nih.gov/nucleotide/87622041>
GI identifier

[FASTA](#) [Graphics](#)

[Display Settings](#): GenBank

Rift Valley fever virus strain 2250/74 segment S, complete sequence

GenBank: DQ380143.1

[Go to](#):

```

LOCUS      DQ380143                1690 bp    RNA        linear    VRL 07-NOV-2007
DEFINITION Rift Valley fever virus strain 2250/74 segment S, complete
            sequence.
ACCESSION  DQ380143
VERSION    DQ380143.1  GI:87622041
KEYWORDS   .
SOURCE     Rift Valley fever virus
ORGANISM   Rift Valley fever virus
            Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
REFERENCE  1 (bases 1 to 1690)
AUTHORS    Bird,B.H., Khristova,M.L., Rollin,P.E., Ksiazek,T.G. and
            Nichol,S.T.
TITLE      Complete genome analysis of 33 ecologically and biologically
            diverse Rift Valley fever virus strains reveals widespread virus
            movement and low genetic diversity due to recent common ancestry
JOURNAL    J. Virol. 81 (6), 2805-2816 (2007)
PUBMED    17192303
REFERENCE  2 (bases 1 to 1690)
AUTHORS    Bird,B.H., Khristova,M.L. and Nichol,S.T.
TITLE      Direct Submission
JOURNAL    Submitted (27-JAN-2006) Special Pathogens Branch, Centers for
            Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA
            30333, USA

FEATURES   Location/Qualifiers
            source                1..1690
                                   /organism="Rift Valley fever virus"
                                   /mol_type="genomic RNA"
                                   /strain="2250/74"
                                   /isolation_source="fetus"
                                   /host="bovine"
                                   /db_xref="taxon:11588"
                                   /segment="S"
                                   /country="Zimbabwe: Beatrice"
                                   /collection_date="1974"

            primer_bind          1..22
            gene                  35..832
    
```

Mesh Countries Recursive [rift valley fever virus](#)

Mesh Country	Organism	NCBI Rank	Sequences	Publications
brazil	belterra virus	no rank	3	0
brazil	icoaraci virus	no rank	3	0
burkina faso	rift valley fever virus	species	4	2
central african republic	rift valley fever virus	species	8	8
chad	rift valley fever virus	species	2	2
egypt	rift valley fever virus	species	26	42
guinea	rift valley fever virus	species	7	0
kenya	rift valley fever virus	species	101	41
madagascar	rift valley fever virus	species	93	7
mauritania	rift valley fever virus	species	10	18
nigeria	rift valley fever virus	species	0	8
saudi arabia	rift valley fever virus	species	2	9
senegal	rift valley fever virus	species	1	22
somalia	rift valley fever virus	species	1	0
south africa	rift valley fever virus	species	11	13
sudan	rift valley fever virus	species	0	10
tanzania	rift valley fever virus	species	0	5
uganda	rift valley fever virus	species	7	0
zimbabwe	rift valley fever virus	species	26	2




Publications

Title	Abstract	Year	PMID
genetic variation among geographic isolates of rift valley fever virus.	the genetic variation of rift valley fever virus (rvfv) was estimated by sequencing a portion of the m segment rna of 22 isolates from a variety of host species collected over 34 years in 6 african countries. the m segment rna of the egyptian isolate, zh501, which has been molecularly cloned and sequenced, was used as a reference for these comparisons. specific gene regions, responsible for antigenic determinants presumed to play a role in protection against disease, were emphasized in these inv ...	1988	2462795
[rift valley fever and phleboviruses in the central african republic].	during 1984 and 1985, six rift valley fever virus strains (rvf) were isolated in central african republic, among them five from human samples. three strains were isolated in 1985 at the end of the rainy season, from sera of patients dead with severe jaundice with haemorrhagic syndrome, what could evoke a little outbreak. at this occasion, these rvf strains and the other strains of phleboviruses previously isolated in car, were antigenically compared by immuno-fluorescent assay (ifa) and compleme ...	1988	3042177
	19 arboviruses pathogenic for human are present in the central african		

Displaying items 1 - 8 of

PubMed

Search

 RSS [Save search](#) [Advanced](#)[Help](#)[Display Settings:](#) Summary, 20 per page, Sorted by Recently Added[Send to:](#) [Filters:](#) [Manage Filters](#)

Results: 1 to 20 of 739

<< First < Prev Page 1 of 37 Next > Last >>

- [Rift Valley fever virus NSs inhibits host transcription independently of the degradation of dsRNA-dependent protein kinase PKR.](#)
1. Kalveram B, Lihoradova O, Indran SV, Lokugamage N, Head JA, Ikegami T.
Virology. 2012 Oct 12. pii: S0042-6822(12)00472-2. doi: 10.1016/j.virol.2012.09.031. [Epub ahead of print]
PMID: 23063407 [PubMed - as supplied by publisher]
[Related citations](#)
- [Acid-activated structural reorganization of the Rift Valley fever virus Gc fusion protein.](#)
2. de Boer SM, Kortekaas J, Spel L, Rottier PJ, Moormann RJ, Bosch BJ.
J Virol. 2012 Oct 3. [Epub ahead of print]
PMID: 23035232 [PubMed - as supplied by publisher]
[Related citations](#)
- [Genetic Subpopulations of Rift Valley Fever ZH548, MP-12 and Recombinant MP-12 Strains.](#)
3. Lokugamage N, Freiberg AN, Morrill JC, Ikegami T.
J Virol. 2012 Oct 3. [Epub ahead of print]
PMID: 23035230 [PubMed - as supplied by publisher]
[Related citations](#)
- [Functional Analysis of Rift Valley Fever Virus NSs Encoding a Partial Truncation.](#)
4. Head JA, Kalveram B, Ikegami T.
PLoS One. 2012;7(9):e45730. doi: 10.1371/journal.pone.0045730. Epub 2012 Sep 19.
PMID: 23029207 [PubMed - in process] [Free PMC Article](#)
[Related citations](#)
- [European ring trial to evaluate ELISAs for the diagnosis of infection with Rift Valley fever virus.](#)
5. Kortekaas J, Kant J, Vloet R, Cêtre-Sossah C, Marianneau P, Lacote S, Banyard AC, Jeffries C, Eiden

[Titles with your search terms](#)

Rift Valley fever virus.

[Curr Mol Med. 2005]

AMP-activated kinase restricts **Rift Valley fever virus** infection by inhibiting f [PLoS Pathog. 2012]

Chemotactic and inflammatory responses in the liver and brain are as [PLoS Negl Trop Dis. 2012]

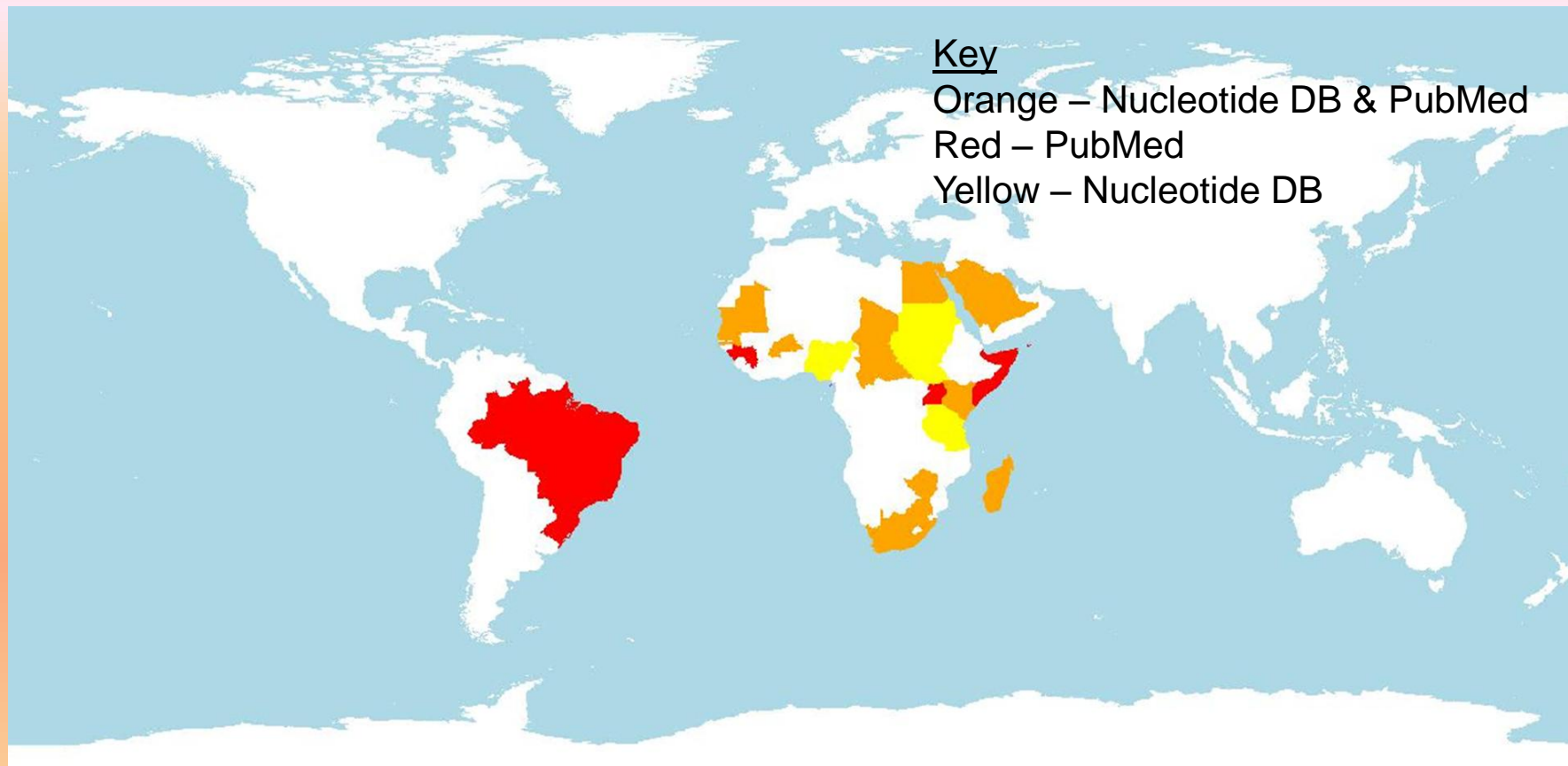
[See more...](#)[208 free full-text articles in PubMed Central](#)

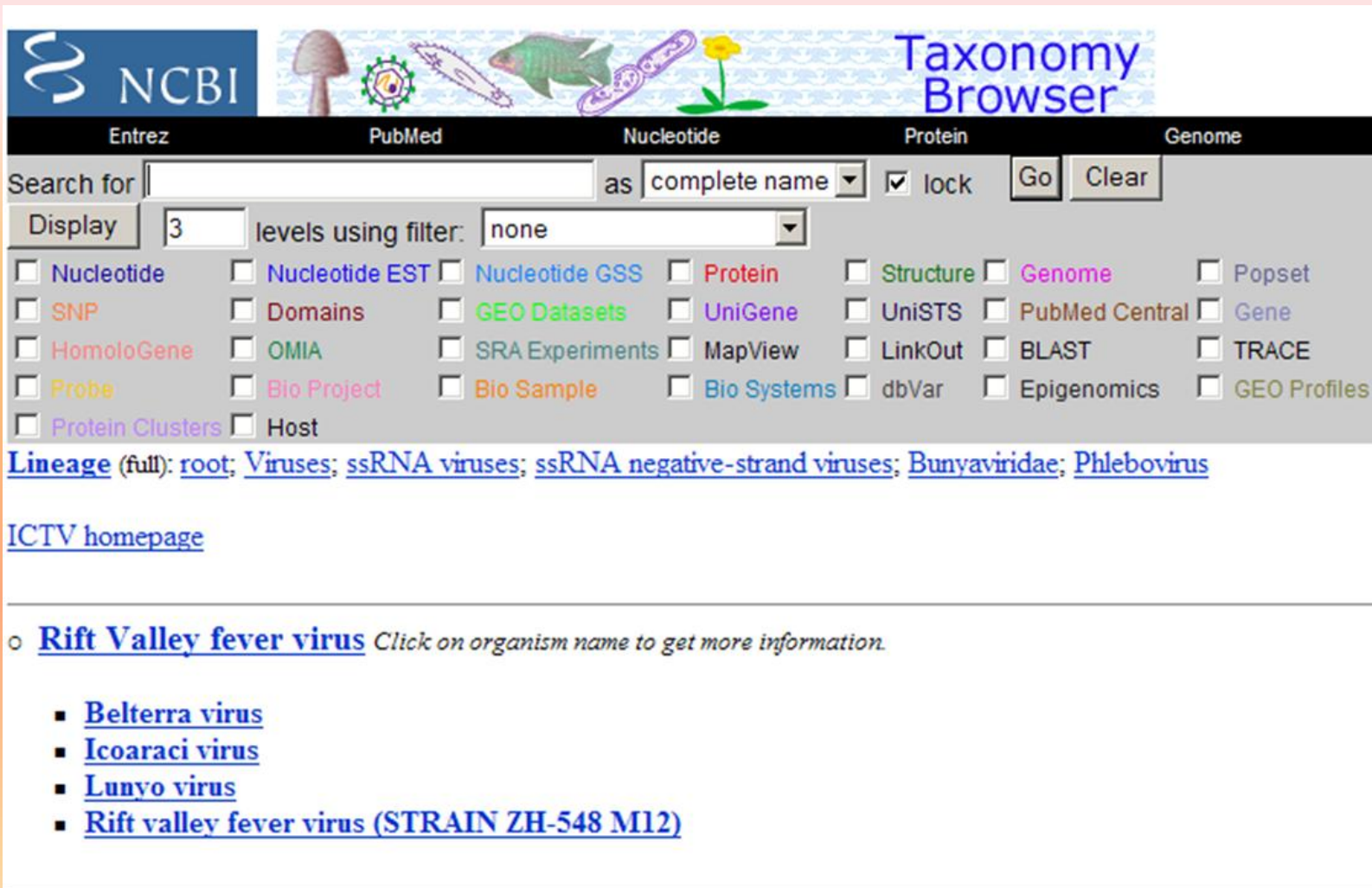
Functional Analysis of Rift Valley Fever Virus NSs Encoding a Partial Truncal [PLoS One. 2012]

Combining hydrology and mosquito population models to identify the [PLoS Negl Trop Dis. 2012]

An epidemiological model of rift valley fever with spatial dynam [Comput Math Methods Med. 2012]

[See all \(208\)...](#)[Find related data](#)Database: [Search details](#)





NCBI Taxonomy Browser

Entrez PubMed Nucleotide Protein Genome

Search for as lock

Display levels using filter:

Nucleotide Nucleotide EST Nucleotide GSS Protein Structure Genome Popset
 SNP Domains GEO Datasets UniGene UniSTS PubMed Central Gene
 HomoloGene OMIA SRA Experiments MapView LinkOut BLAST TRACE
 Probe Bio Project Bio Sample Bio Systems dbVar Epigenomics GEO Profiles
 Protein Clusters Host

Lineage (full): [root](#); [Viruses](#); [ssRNA viruses](#); [ssRNA negative-strand viruses](#); [Bunyaviridae](#); [Phlebovirus](#)

[ICTV homepage](#)

o [Rift Valley fever virus](#) *Click on organism name to get more information.*

- [Belterra virus](#)
- [Icoaraci virus](#)
- [Lunyo virus](#)
- [Rift valley fever virus \(STRAIN ZH-548 M12\)](#)

[Display Settings:](#) GenBank

Belterra virus strain Be An 356637 N protein gene, partial cds

GenBank: EF076013.1

[FASTA](#) [Graphics](#)

[Go to:](#)

```

LOCUS           EF076013                463 bp    cRNA    linear    VRL 16-JUN-2007
DEFINITION     Belterra virus strain Be An 356637 N protein gene, partial cds.
ACCESSION     EF076013
VERSION       EF076013.1  GI:134254649
KEYWORDS      .
SOURCE        Belterra virus
  ORGANISM    Belterra virus
              Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
REFERENCE     1 (bases 1 to 463)
  AUTHORS     Xu,F., Liu,D., Nunes,M.R., DA Rosa,A.P., Tesh,R.B. and Xiao,S.Y.
  TITLE      Antigenic and genetic relationships among Rift Valley fever virus
              and other selected members of the genus Phlebovirus (Bunyaviridae)
  JOURNAL     Am. J. Trop. Med. Hyg. 76 (6), 1194-1200 (2007)
  PUBMED     17556635
REFERENCE     2 (bases 1 to 463)
  AUTHORS     Xiao,S.-Y., Xu,F., Liu,D. and Tesh,R.B.
  TITLE      Direct Submission
  JOURNAL     Submitted (18-OCT-2006) Pathology, University of Texas Medical
              Branch, 301 University Blvd., Galveston, TX 77555, USA
FEATURES      Location/Qualifiers
  source      1..463
              /organism="Belterra virus"
              /mol_type="viral cDNA"
              /strain="Be An 356637"
              /db_xref="taxon:426789"
              /segment="3"
              /country="Brazil"
  CDS         35..>463
              /codon_start=1
              /product="N protein"
    
```

		true location status		Total
		location correct (+)	location incorrect (-)	
Location status using MeSH term	MeSH term for location (+)	a	b	a+b
	No MeSH term for location (-)	c	d	c+d
Total		a+c	b+d	a+b+c+d

Test sensitivity = $a/(a+c)$

Test specificity = $d/(b+d)$

Predictive value of a positive test result = $a/(a+b)$

True prevalence = $(a+c)/(a+b+c+d)$

Apparent prevalence = $(a+b)/(a+b+c+d)$

Predictive value of positive test = 0.95 (SE=0.05)

Likelihood of relationship between pathogen + country MeSH term not affected by:

- pathogenic status ($P=0.393$)
- taxonomic division ($P>0.05$).

Threshold of 5 papers

Likelihood of the relationship being false = $(1 - 0.95)^5 = 0.0000003$
 i.e. 99.99...% certainty

Binomial regression modelling approach

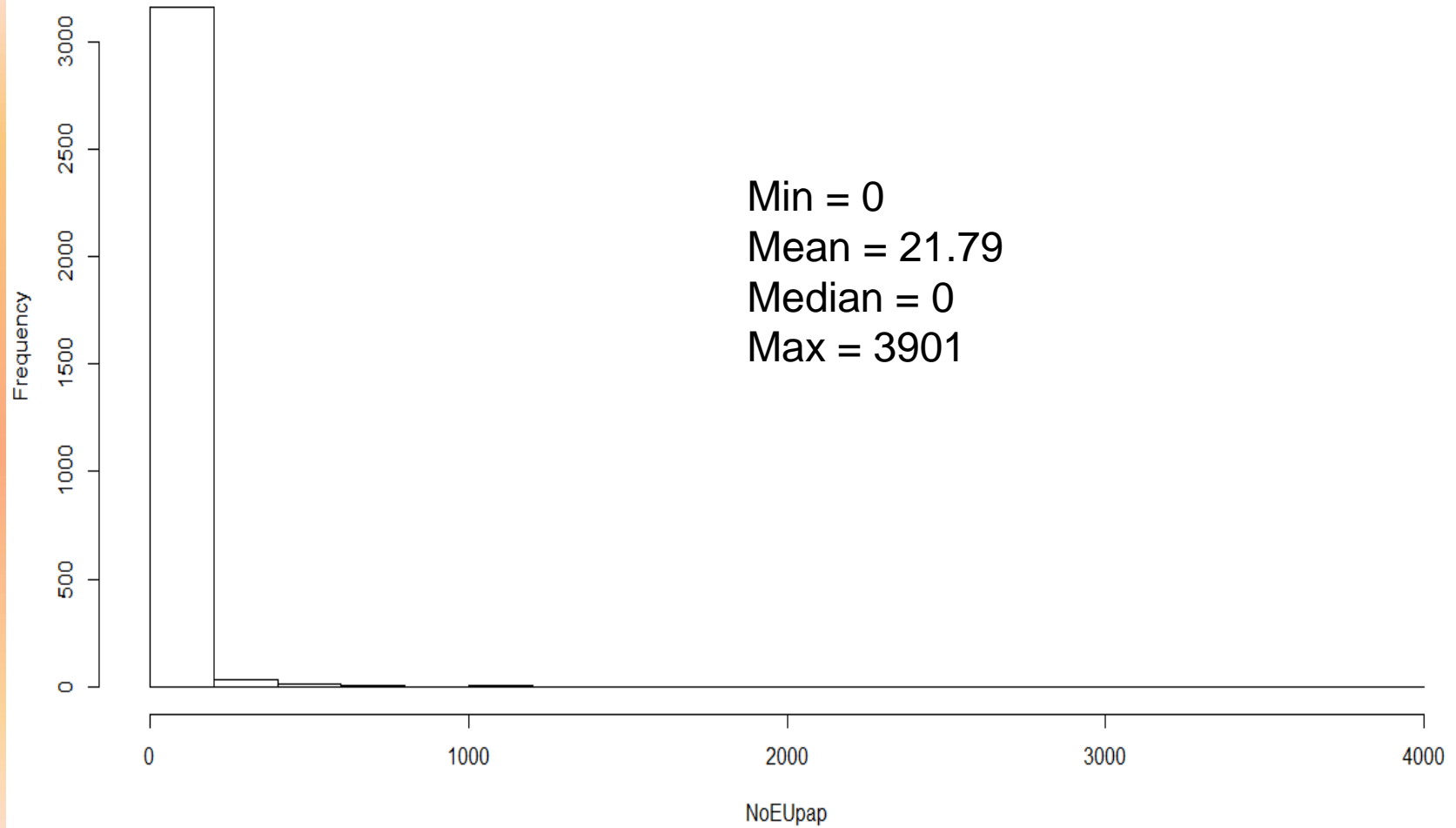
Probability of at least one Nucleotide sequence ~ No. PubMed papers

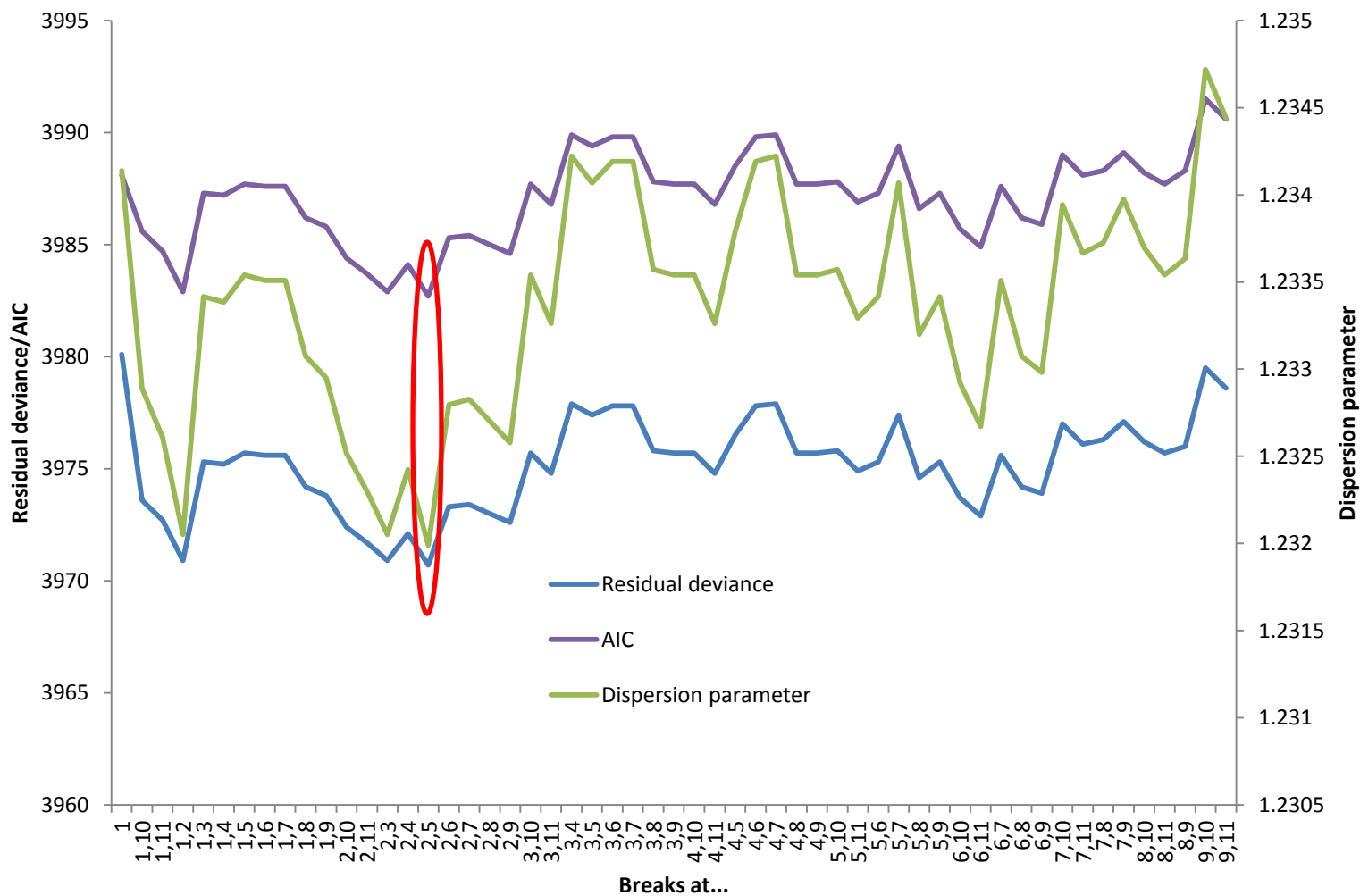
Investigate:

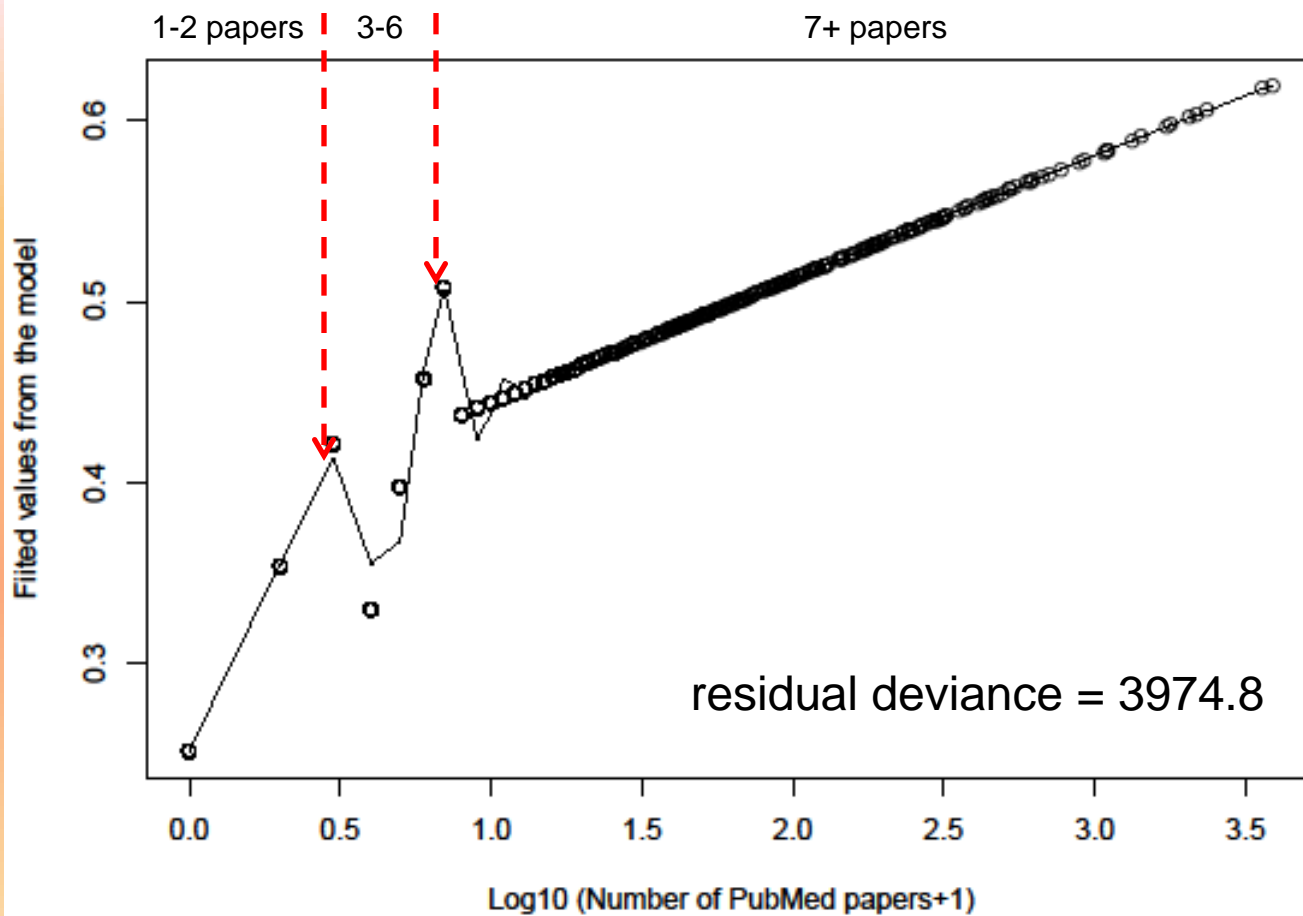
- Quadratic terms
- $\text{Log}_{10}(n+1)$ transformation of x

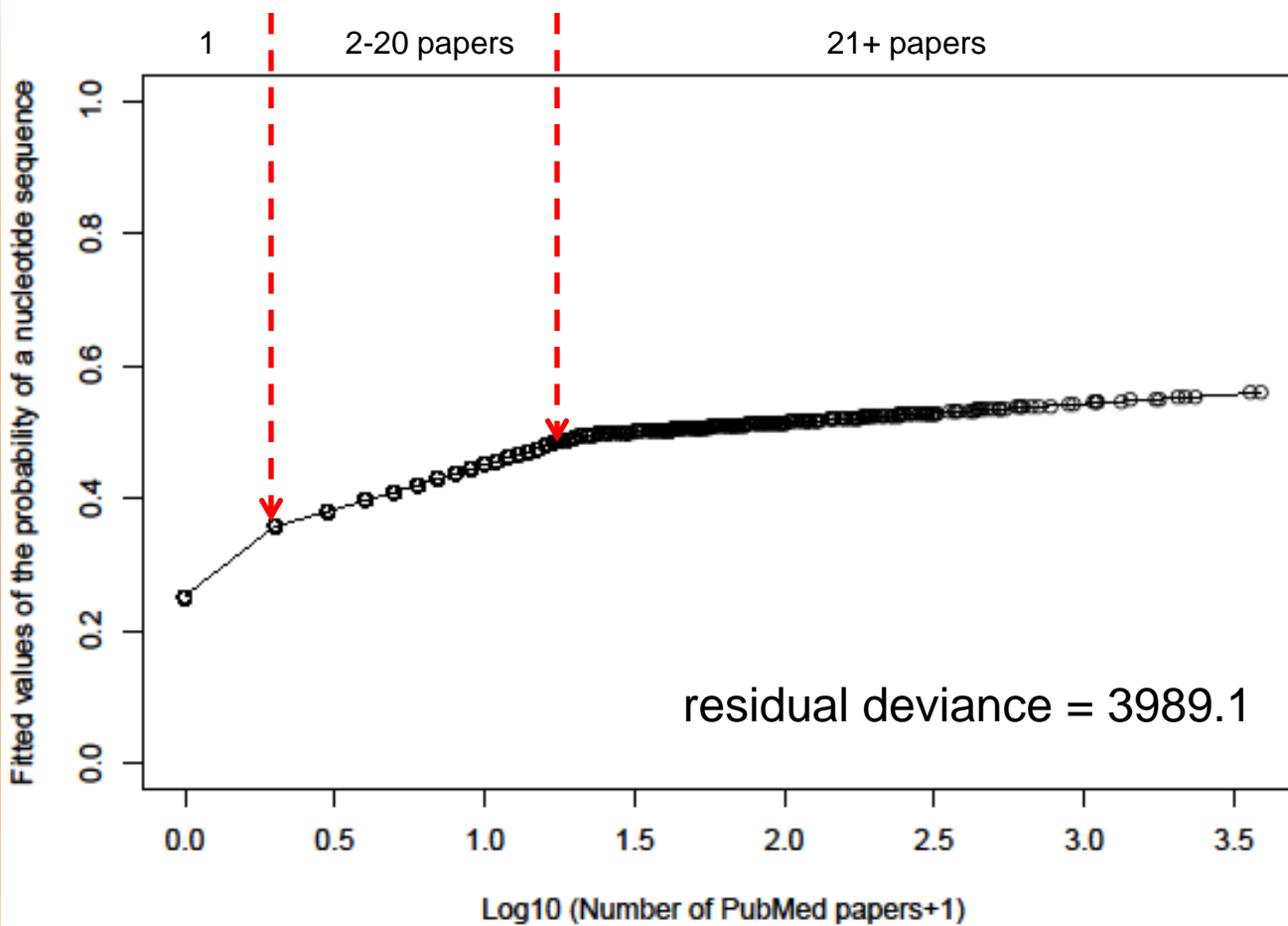
namekey	Nucevid	NoEUpap
1	1	1
2	0	0
3	1	0
4	0	5
5	0	1
6	1	0
7	0	1
8	1	0
9	0	4
10	0	0
11	0	0
12	0	0
13	0	0
14	1	0
15	0	5
16	1	1
17	0	15
18	1	2
19	1	301
20	1	35
21	1	2
22	1	1
23	1	7
24	0	8
25	0	1
26	1	0
27	0	1
28	0	0
29	0	0

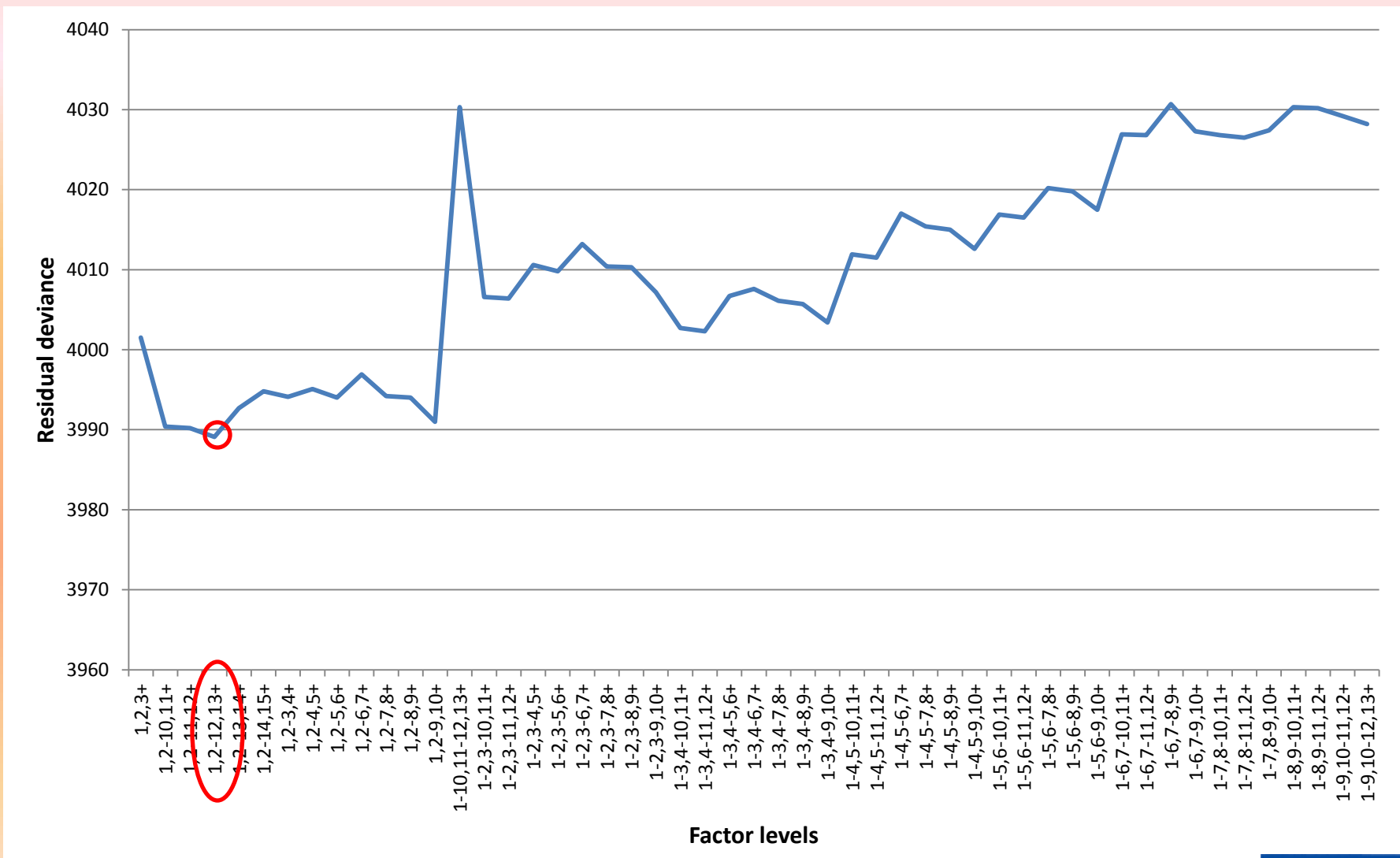
Histogram of NoEUpap











Probability of Nucleotide sequence ~ No. PubMed papers (factor), 2 breaks

No. of papers	Odds ratio	Confidence intervals
1 (Baseline)	-	-
2-12	1.86	1.55-2.22
13+	2.89	2.35-3.55

In summary

- Pathogen-host interaction data incorporated into EID2 using NCBI Nucleotide (Nucleotide) database metadata
- Location information for pathogens incorporated using Nucleotide metadata
- Location information also included using PubMed searches
- Threshold of 5 papers in which pathogen name + country MeSH term had 99.9% certainty of 'true' relationship
- Binomial GLMs also suggested 2+ papers was good threshold, though ideal number of papers changed dependent upon model – essentially the more the better!

Acknowledgements



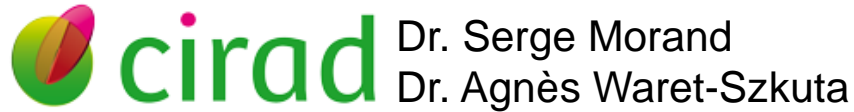
Department of Geography



Prof. Matthew Baylis
(Dr. Phil Hepworth)
Dr. K. Marie McIntyre

Dr. Cyril Caminade
Mr. Andrew McCaldon
Mr. Peris Roberts
Dr. Andy Morse

(Dr. Christian Setzkorn)



Maastricht University

Prof. Pim Martens
Dr. Maud Huynens
Ms. Su-Mia Akin



Ms. Emma Bennett
Dr. Steve Leach
Mr. Jolyon Medlock
Prof. John R. Stephenson



Dr. Helen Roberts
Dr. Scott Sellers
Dr. Peter Stevenson

