# Optimized design of broadly detecting qPCR Primers and Probes using a conservation and hybridization prediction algorithm.

Jonas Blomberg, MD PhD,
professor emeritus,
clinical virology,
Department of
Medical Sciences,
Uppsala University

# The problem

## 1. Target sequence variation

-Number of viruses increased much during the last 10 years

-Viruses lack highly conserved stretches like 16S rRNA sequences

-Broad detection of several viral species requires special tricks

-If the viruses do not offer conserved targets, primers and probes must somehow encompass target variation, yet preserve specificity.

Thus,

-Variable target sequences require **variation tolerant primers and probes**

# The problem

## 2. Need for multiplexing

-Syndromic qPCR panels may have to cover 20-30 viral species

-Not practical to have more than 5 fluorophores per amplification well

-Too many oligos per well give interference phenomena

# The problem

## 3. Avoidance of false positives (contamination)

-Best is to amplify within a closed container.
Opening after amplification invites trouble.

# The problem

## 4. Limits to primer and probe degeneration

-Number of variants increase for each degeneration, diminishes sensitivity

-”Universal” bases may or may not work

# Our solution: ConSort, and VOCMA

A new hybridization algorithm (NucZip)
handles both degeneration and
a universal base (inosine)

Long primers and probes (>30nt) are more
variation tolerant than short ones.
Nuczip handles that.

# **Other solution**s (list is not comprehensive)

**Primer 3**

GC content,
Nearest Neighbor
calculation

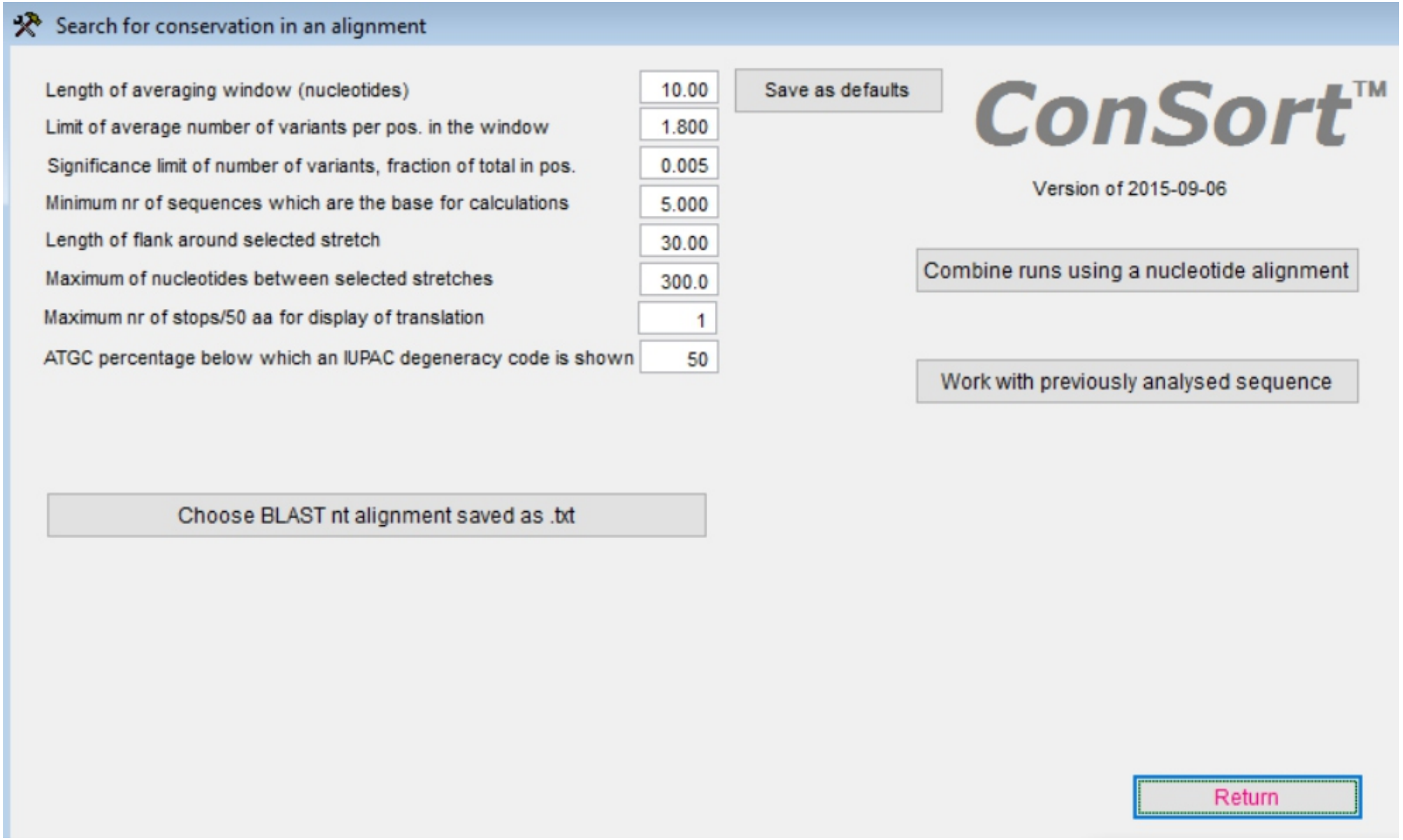
**Netprimer**

Approximation of Tm

**Oligo Analyzer**

**PerlPrimer**

# Large scale automated search for conservation and enhancement of primer and probe design



**ConSort:**

Can handle BLASTn alignments of up to 10 000 sequences

Displays conservation,
from broad overview to detailed scrutiny

Can divide variability into haplotypes

Can test suggested primer/probes against the alignment

Can use NucZip and NN algorithms

Can handle all kinds of degeneration and one ”universal” base (inosine)

Can find primer and probe targets in alignment

Is written in Visual Foxpro version 9 by JB,

an object oriented relational database system

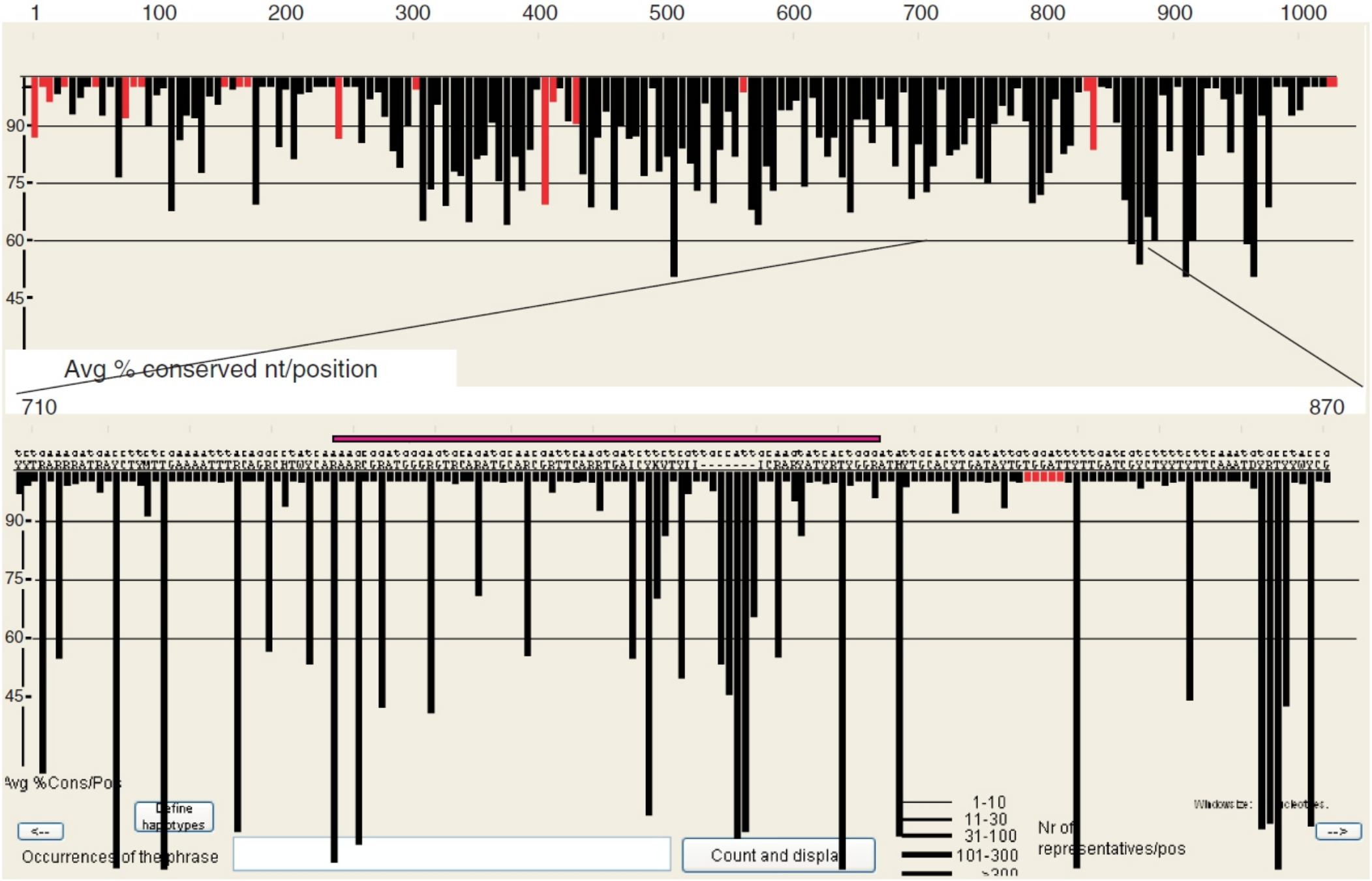
**16 publications:**

Retroviruses,
Influenza,
Coronaviruses,
Hepeviruses,
Rotavirus,
Calicivirus

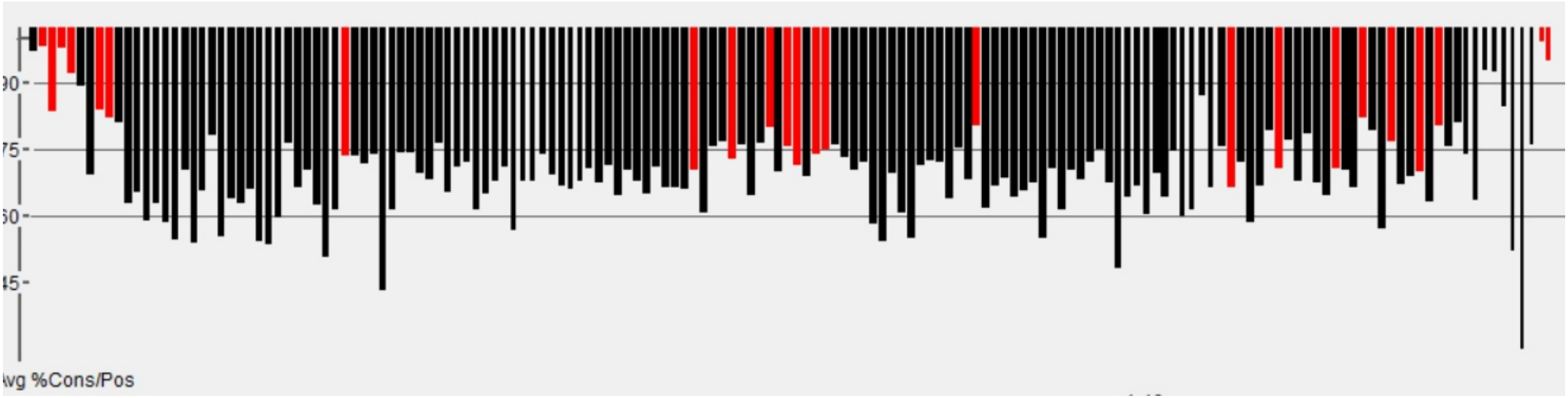
**Unpublished:**

Herpesviruses,
Anelloviruses,
Polyomaviruses,
Parvoviruses,
Flaviviruses

# Task 1: Find conservation




7333 Influenza A segment 7 sequences




Hepatitis C variation

# Task 2: Construct primers and probes

**All more advanced primer and probe analysis programs available are based on**

## Nearest Neighbour model

Hybridization of nucleotides are affected by
- the neigbouring match/mismatch and
- by which type of nt it has as a neighbour

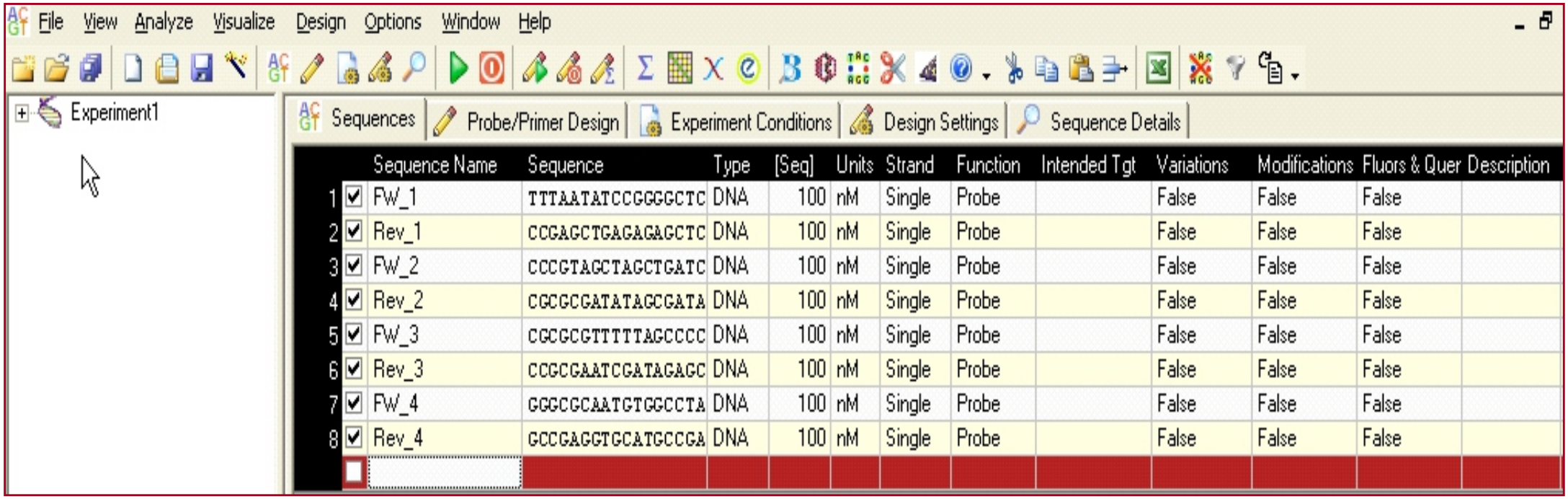
**(OligoAnalyser (IDT), VisualOmp (DNASoftware), mFold, OligoArray …)**

©Christina Öhrmalm

# Multiplex in silico testing

Visual OMP™ Nucleic Acid Software

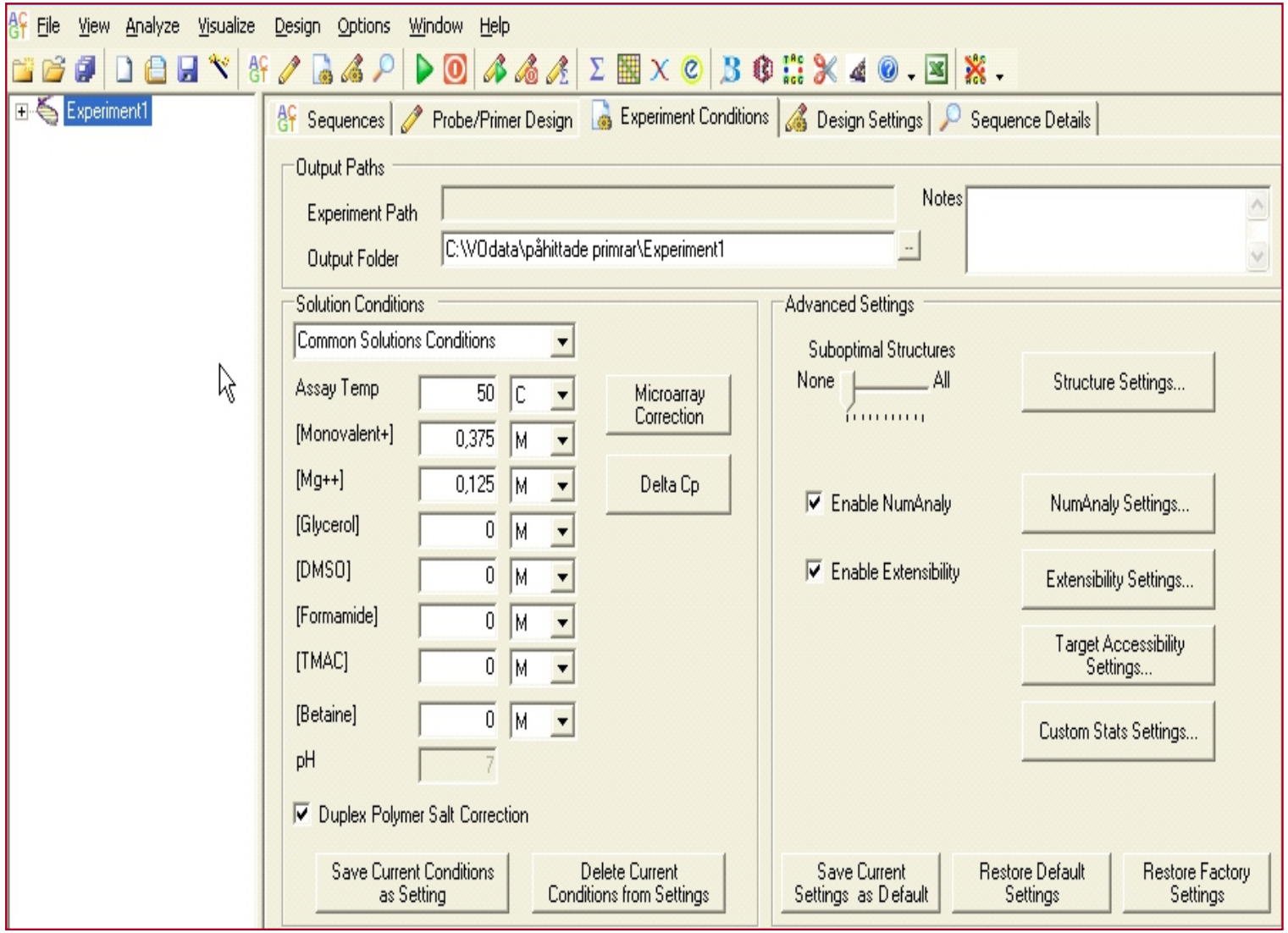
DNA software




| | | Sequence Name | Sequence | Type | [Seq] | Units | Strand | Function | Intended Tgt | Variations | Modifications | Fluors & Quer | Description |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ☑ | FW_1 | TTTAATATCCGGGGCTC | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 2 | ☑ | Rev_1 | CCGAGCTGAGAGAGCTC | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 3 | ☑ | FW_2 | CCCGTAGCTAGCTGATC | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 4 | ☑ | Rev_2 | CGCGCGATATAGCGATA | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 5 | ☑ | FW_3 | CGCGCGTTTTTAGCCCC | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 6 | ☑ | Rev_3 | CCGCGAATCGATAGAGC | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 7 | ☑ | FW_4 | GGGCGCAATGTGGCCTA | DNA | 100 | nM | Single | Probe | | False | False | False | |
| 8 | ☑ | Rev_4 | GCCGAGGTGCATGCCGA | DNA | 100 | nM | Single | Probe | | False | False | False | |
| | ☐ | | | | | | | | | | | | |

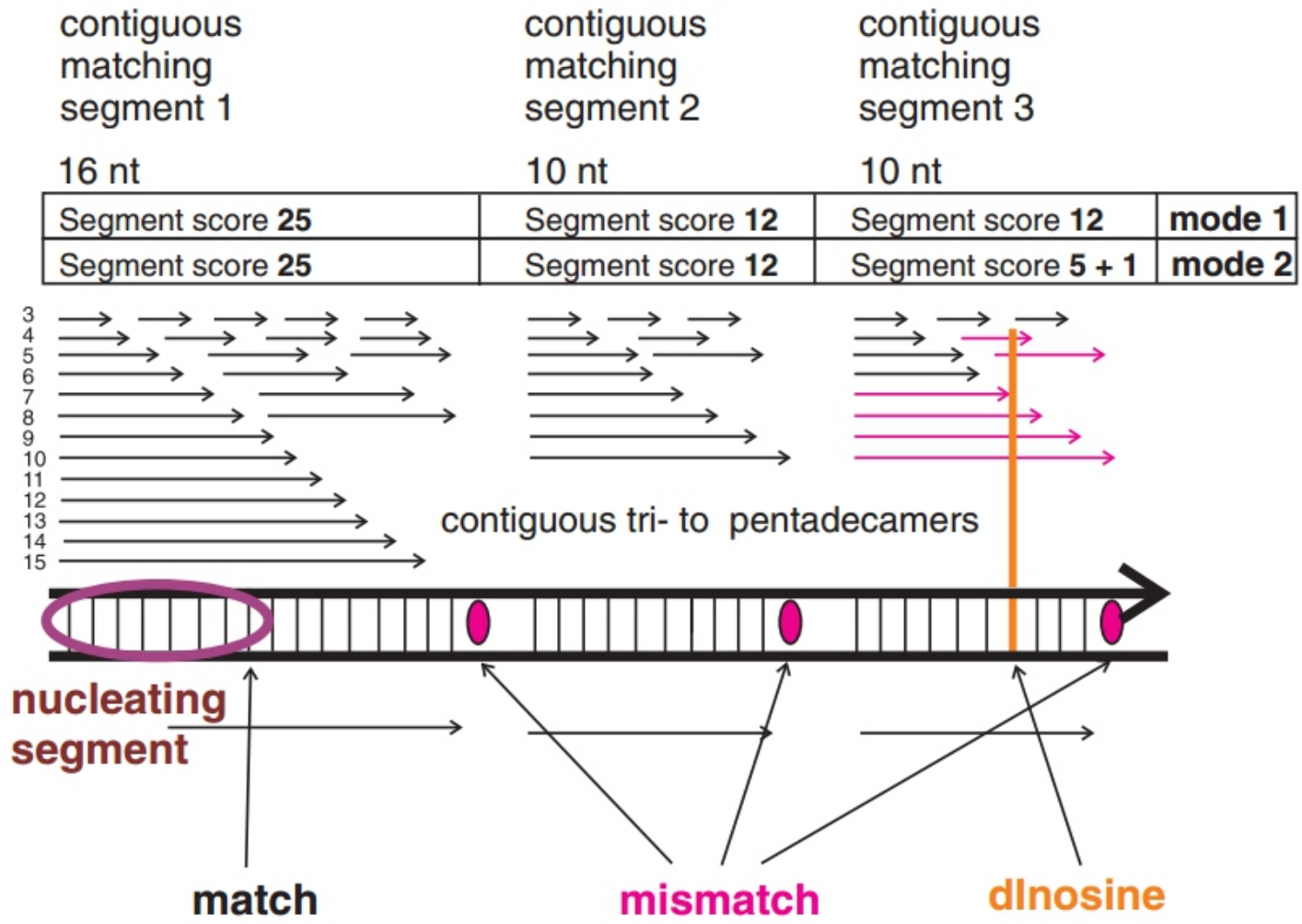
# Visual OMP™ Nucleic Acid Software

DNA software




Set the ”Experiment Conditions”

- Buffer conditions
- Annealing temperature
- No of Suboptimal Structures




Total ZipScore$_{mode1}$: 25+12+12=49

Total ZipScore$_{mode2}$: 25+12+5+1=43

ZipScore$_{downstream}$: 49-dInosinefactor*(49-43)+(dInoCnr*dInoCfactor)

**Figure 8.** The zipping component of NucZip. When a 6–9-bp sequence

How NucZip works

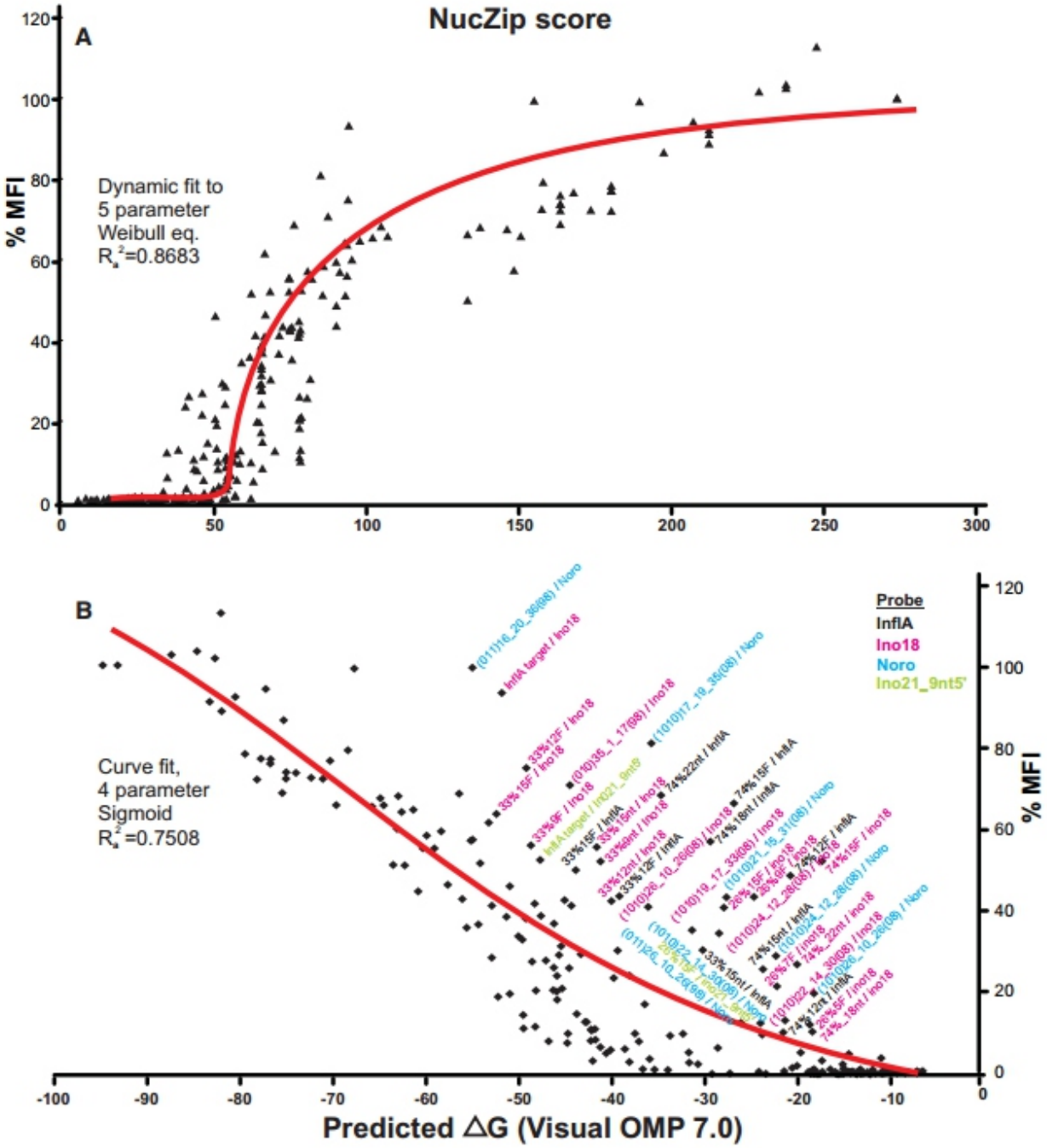



**Figure 6.** Prediction of hybridization using two algorithms. (**A**) NucZip score versus %MFI. The results of the NucZip scoring system versus the %MFI for the same target and probe combinations that were used for Figure 5A. A five parameter Weibull function gave the highest correlation

Our NucZip algorithm was better than the nearest neighbor algorithm for long probes.

# The NucZip Algorithm

```
FUNCTION nuczip
 LPARAMETERS _Target, _Probe, _Linfactor, _Expfactor, _Inosinefactor,;
  _InoCfactor
 _Mers=" tri tetra penta hexa hepta okta nona deka undeka dodeka trideka
  tetradeka pentadeka "
 _cProbe=complement(_probe)
 _Nuc=nucleation(_target,_cprobe)
 _Nuc1points=VAL(SUBSTR(_Nuc,AT("!",_Nuc,3)+1))
 _Nuc2points=VAL(SUBSTR(_Nuc,AT("!",_Nuc,7)+1))
 IF _Nuc1points>_Nuc2points
  _Nucpoints=_Nuc1points
 ELSE
  _Nucpoints=_Nuc2points
 ENDIF
 IF _Nucpoints<5
  RETURN 0
 ENDIF
 SELECT 1
 GOTO top
 _Out=stretchcount(_target, _probe)
 _OutM=","+SUBSTR(_Out,1,AT("_",_Out)-1)
 _OutI=","+SUBSTR(_Out,AT("_",_Out)+1)
 _trim=VAL(SUBSTR(_OutM,AT(",",_OutM,1)+1))
 _tetraM=VAL(SUBSTR(_OutM,AT(",",_OutM,2)+1))
 _pentaM=VAL(SUBSTR(_OutM,AT(",",_OutM,3)+1))
 _hexaM=VAL(SUBSTR(_OutM,AT(",",_OutM,4)+1))
 _heptaM=VAL(SUBSTR(_OutM,AT(",",_OutM,5)+1))
 _oktaM=VAL(SUBSTR(_OutM,AT(",",_OutM,6)+1))
 _nonaM=VAL(SUBSTR(_OutM,AT(",",_OutM,7)+1))
 _dekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,8)+1))
 _undekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,9)+1))
 _dodekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,10)+1))
 _tridekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,11)+1))
 _tetradekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,12)+1))
 _pentadekaM=VAL(SUBSTR(_OutM,AT(",",_OutM,13)+1))
 _triI=VAL(SUBSTR(_OutI,AT(",",_OutI,1)+1))
 _tetraI=VAL(SUBSTR(_OutI,AT(",",_OutI,2)+1))
 _pentaI=VAL(SUBSTR(_OutI,AT(",",_OutI,3)+1))
 _hexaI=VAL(SUBSTR(_OutI,AT(",",_OutI,4)+1))
 _heptaI=VAL(SUBSTR(_OutI,AT(",",_OutI,5)+1))
 _oktaI=VAL(SUBSTR(_OutI,AT(",",_OutI,6)+1))
 _nonaI=VAL(SUBSTR(_OutI,AT(",",_OutI,7)+1))
 _dekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,8)+1))
 _undekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,9)+1))
 _dodekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,10)+1))
 _tridekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,11)+1))
 _tetradekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,12)+1))
 _pentadekaI=VAL(SUBSTR(_OutI,AT(",",_OutI,13)+1))
```

NAR 2010

Together with Christina Öhrmalm, i developed a new hybridization estimation system, using a program written in Visual FoxPro

*Published online 22 September 2010*

*Nucleic Acids Research, 2010, Vol. 38, No. 21* ***e195***
*doi:10.1093/nar/gkq777*

# Hybridization properties of long nucleic acid probes for detection of variable target sequences, and development of a hybridization prediction algorithm

**Christina Öhrmalm[1], Magnus Jobs[1,2], Ronnie Eriksson[1], Sultan Golbob[1], Amal Elfaitouri[1], Farid Benachenhou[1], Maria Strømme[3] and Jonas Blomberg[1,*]**

[1]Clinical Virology, Department of Medical Sciences, Uppsala University and Academic Hospital, 751 85 Uppsala, [2]School of Health and Social Sciences, Högskolan Dalarna, Falun and [3]Nanotechnology and Functional Materials, The Ångström Laboratory, Uppsala University, Box 534, 75121 Uppsala, Sweden

Received February 13, 2010; Revised August 10, 2010; Accepted August 15, 2010

# Use of NucZip to predict PCR amplifiability

Lo-Alter outer portion of nested gag primers

A

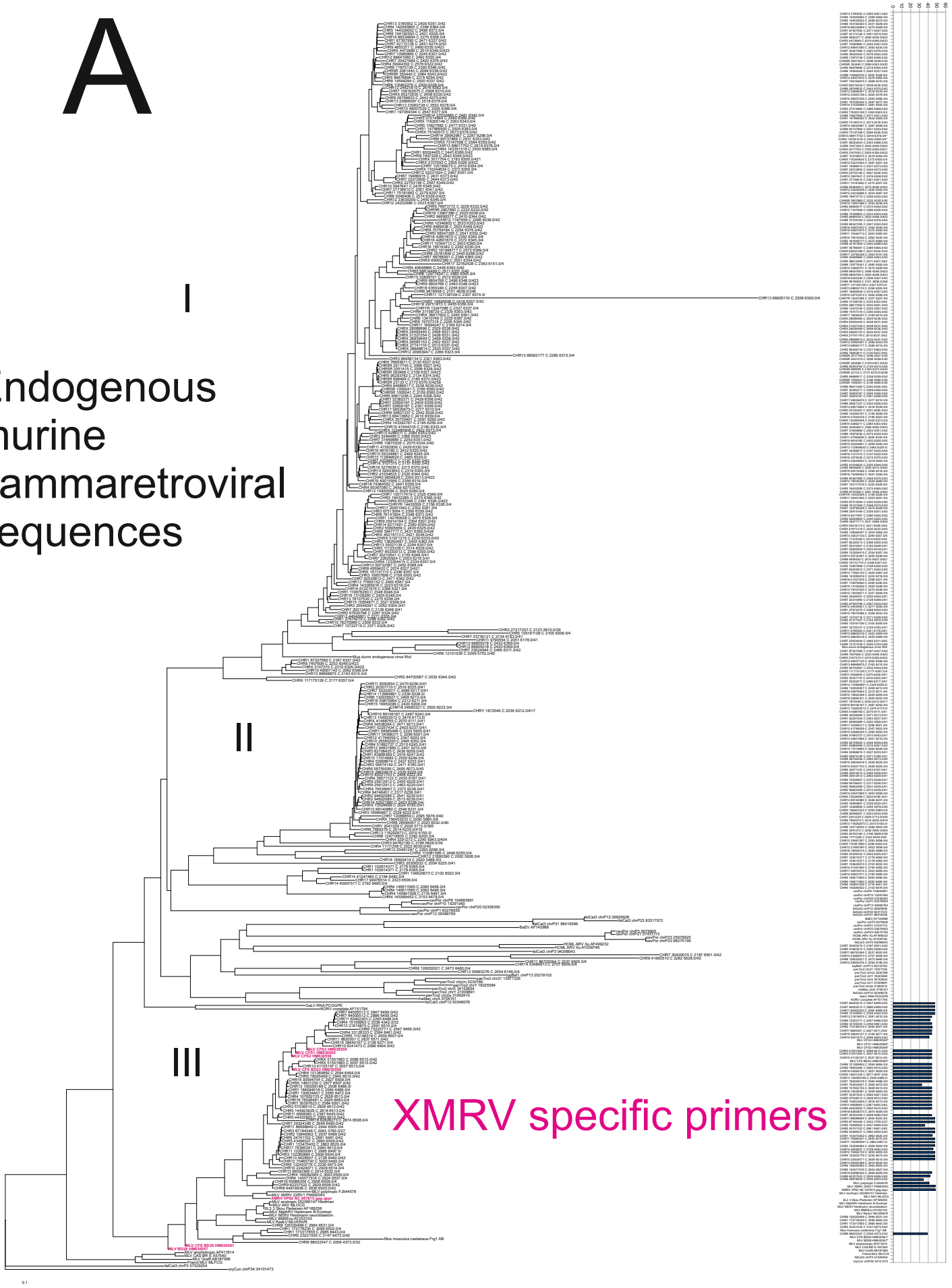
I

Endogenous murine gammaretroviral sequences

II

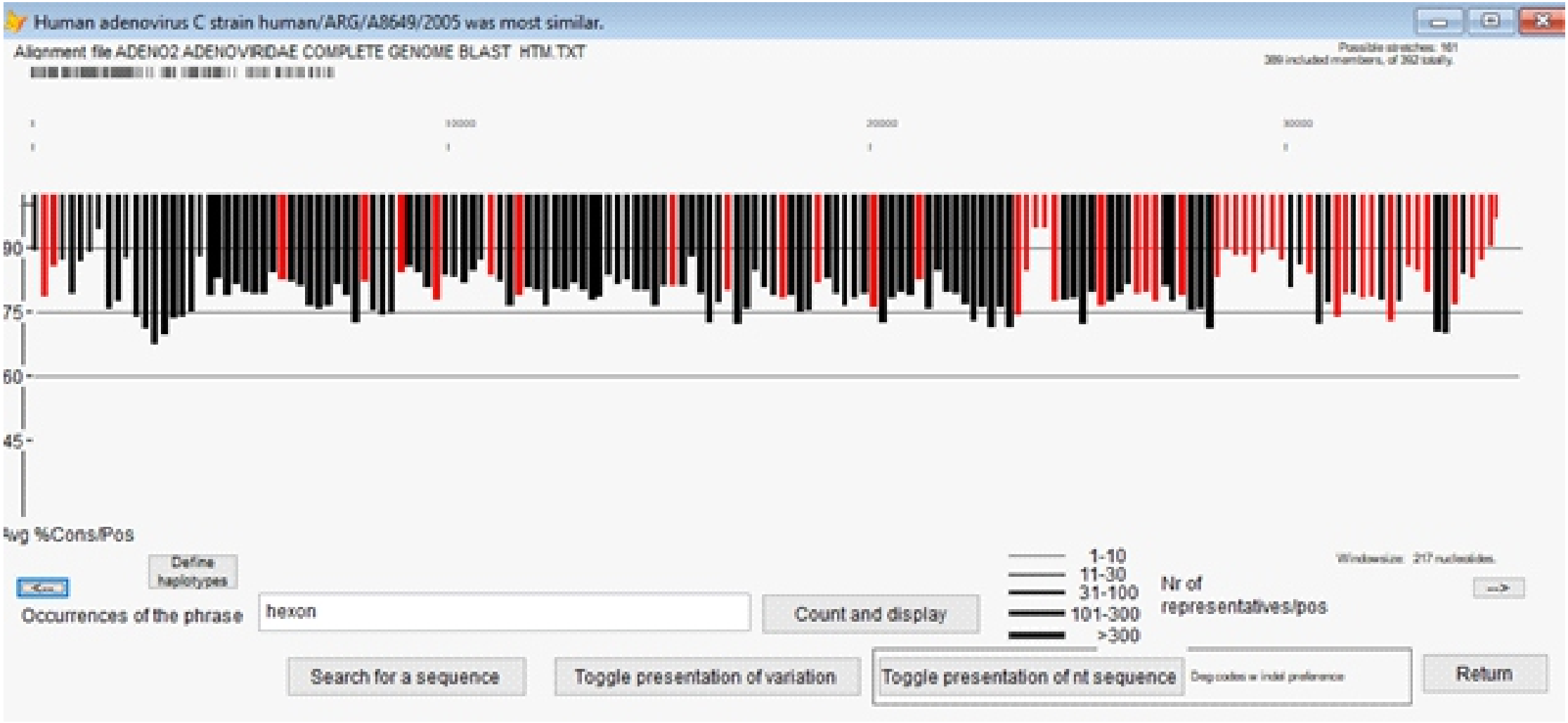
III

XMRV specific primers

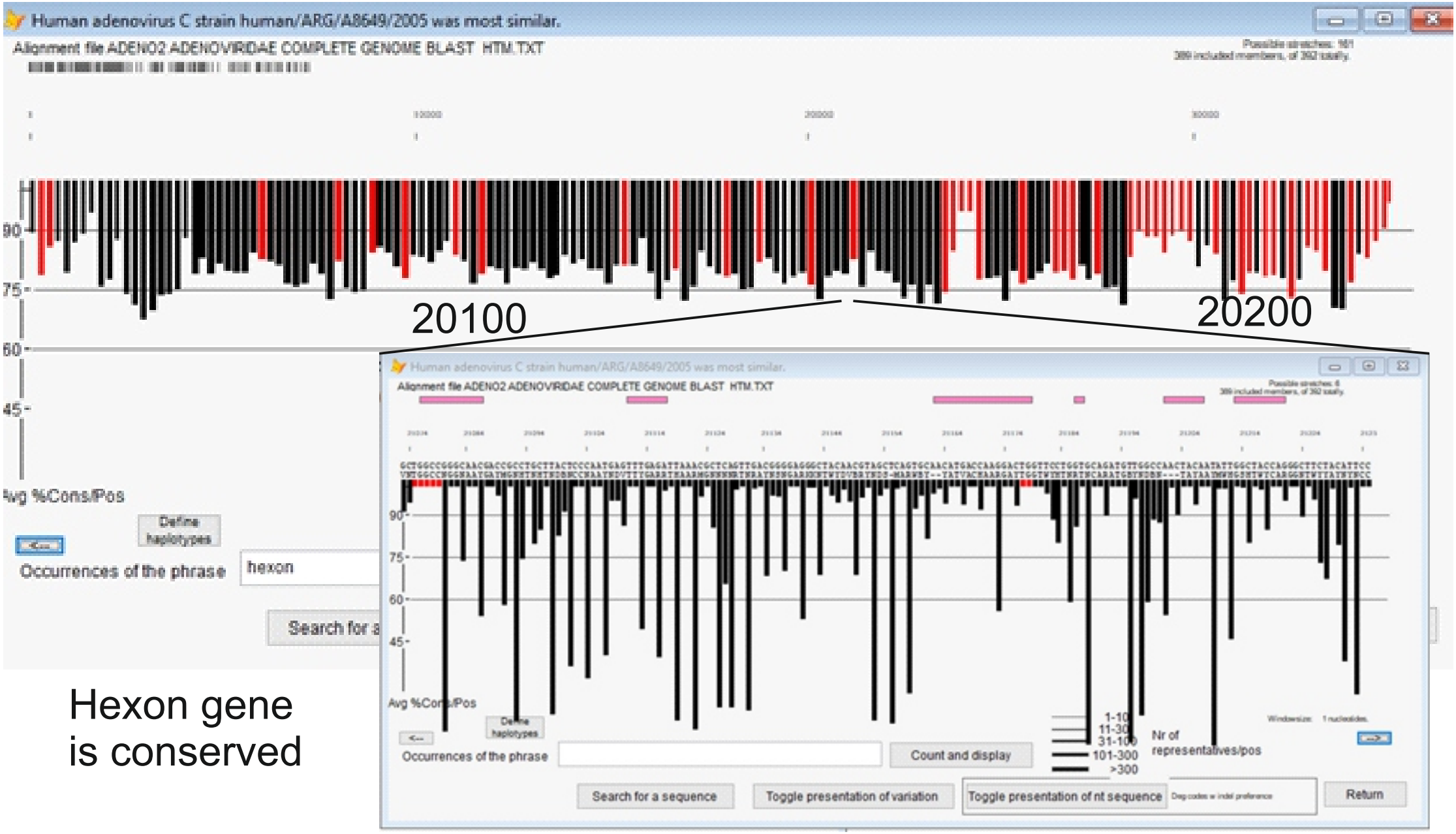


Elfaitouri et al. Plos One Oct 2011. Program written by JB.

# Initial representation of variability, Adeno type 2 alignment



# Initial representation of variability, Adeno type 2 alignment



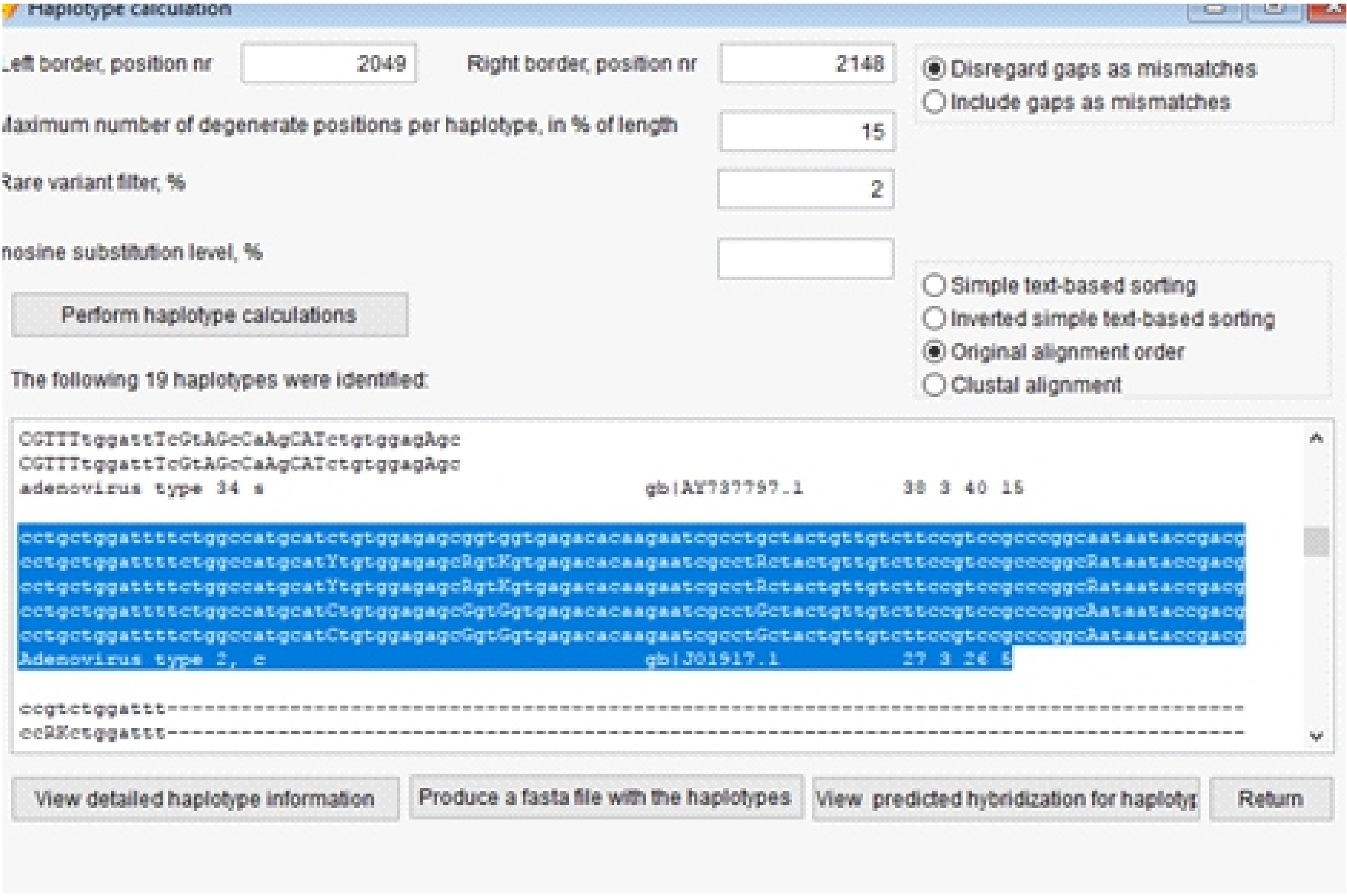

Hexon gene is conserved

# Another way of looking at variability: Haplotypes

Haplotypes are members of a sequence set which share a predetermined similarity.

Ideally, members within a haplotype can be represented by one sequence, degenerated or not.

Haplotypes allow construction of **primer/probe ”swarms”** which can cooperate in amplifying a large and complex set of target sequences.


Haplotype calculation
Left border, position nr
2049
Right border, position nr
2148
Disregard gaps as mismatches
Include gaps as mismatches
Maximum number of degenerate positions per haplotype, in % of length
15
Rare variant filter, %
2
nosine substitution level, %
Simple text-based sorting
Inverted simple text-based sorting
Original alignment order
Clustal alignment
Perform haplotype calculations
The following 19 haplotypes were identified:
CGTTTtggatttcGtAGcCaAgCATctgtggagAgc
CGTTTtggatttcGtAGcCaAgCATctgtggagAgc
adenovirus type 34 s gb|AY737797.1 38 3 40 15
cctgctggatttttctggccatgcatctgtggagagcggtggtgagacacaagaatcgcctgctactgttgtcttccgtccgcccggcaataataccgacg
cctgctggatttttctggccatgcatYtgtggagagcRgtKgtgagacacaagaatcgcctRctactgttgtcttccgtccgcccggcRataataccgacg
cctgctggatttttctggccatgcatYtgtggagagcRgtKgtgagacacaagaatcgcctRctactgttgtcttccgtccgcccggcRataataccgacg
cctgctggatttttctggccatgcatCtgtggagagcGgtGgtgagacacaagaatcgcctGctactgttgtcttccgtccgcccggcAataataccgacg
cctgctggatttttctggccatgcatCtgtggagagcGgtGgtgagacacaagaatcgcctGctactgttgtcttccgtccgcccggcAataataccgacg
Adenovirus type 2, c gb|J01917.1 27 3 26 5
ccgtctggattt----------
ccRRctggattt----------
View detailed haplotype information
Produce a fasta file with the haplotypes
View predicted hybridization for haplotyp
Return

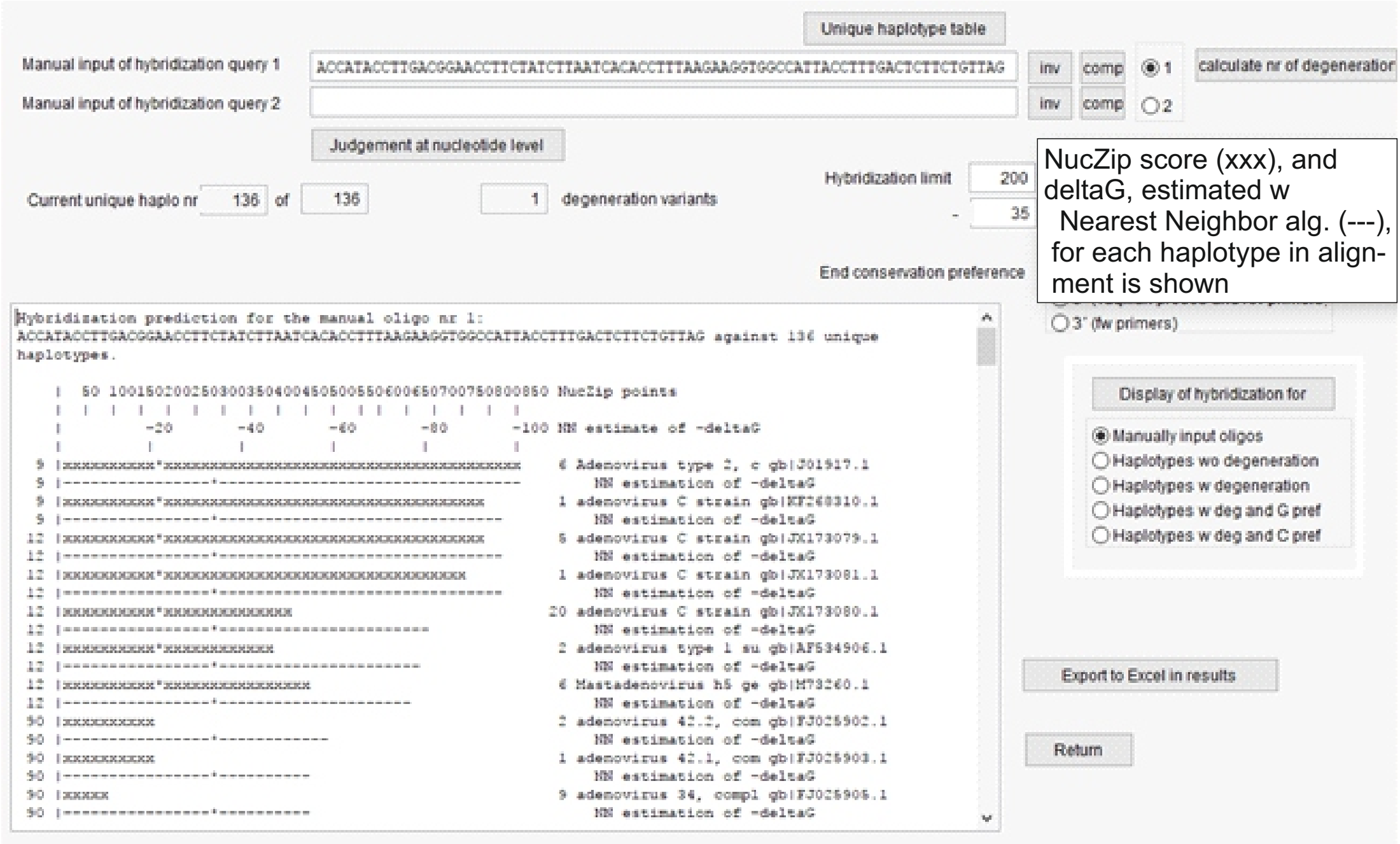
**One of the 19 haplotypes:**

**Consensus** cctgctggattttctggccatgcatctgtggagagcggtggtgagacacaagaatcgcctgctactgttgtcttccgtccgcccggcaataataccgacg
**wIUPAC** cctgctggattttctggccatgcat **Y**tgtggagagc **R**gt**K**gtgagacacaagaatcgcct **R**ctactgttgtcttccgtccgcccggc **R**ataataccgacg
**wIUPAC&gap** cctgctggattttctggccatgcat **Y**tgtggagagc **R**gt**K**gtgagacacaagaatcgcct **R**ctactgttgtcttccgtccgcccggc **R**ataataccgacg
**wGpref&I** cctgctggattttctggccatgcatCtgtggagagcGgtGgtgagacacaagaatcgcctGctactgttgtcttccgtccgcccggcAataataccgacg
**wCpref&I** cctgctggattttctggccat gcatCtgtggagagcGgtGgtgagacacaagaatcgcctGctactgttgtcttccgtccgcccggcAataataccgacg

Haplo ID: Adenovirus type 2, c gb|J01917.1 nr 27 overlaps 3 members 26 deg positions 5

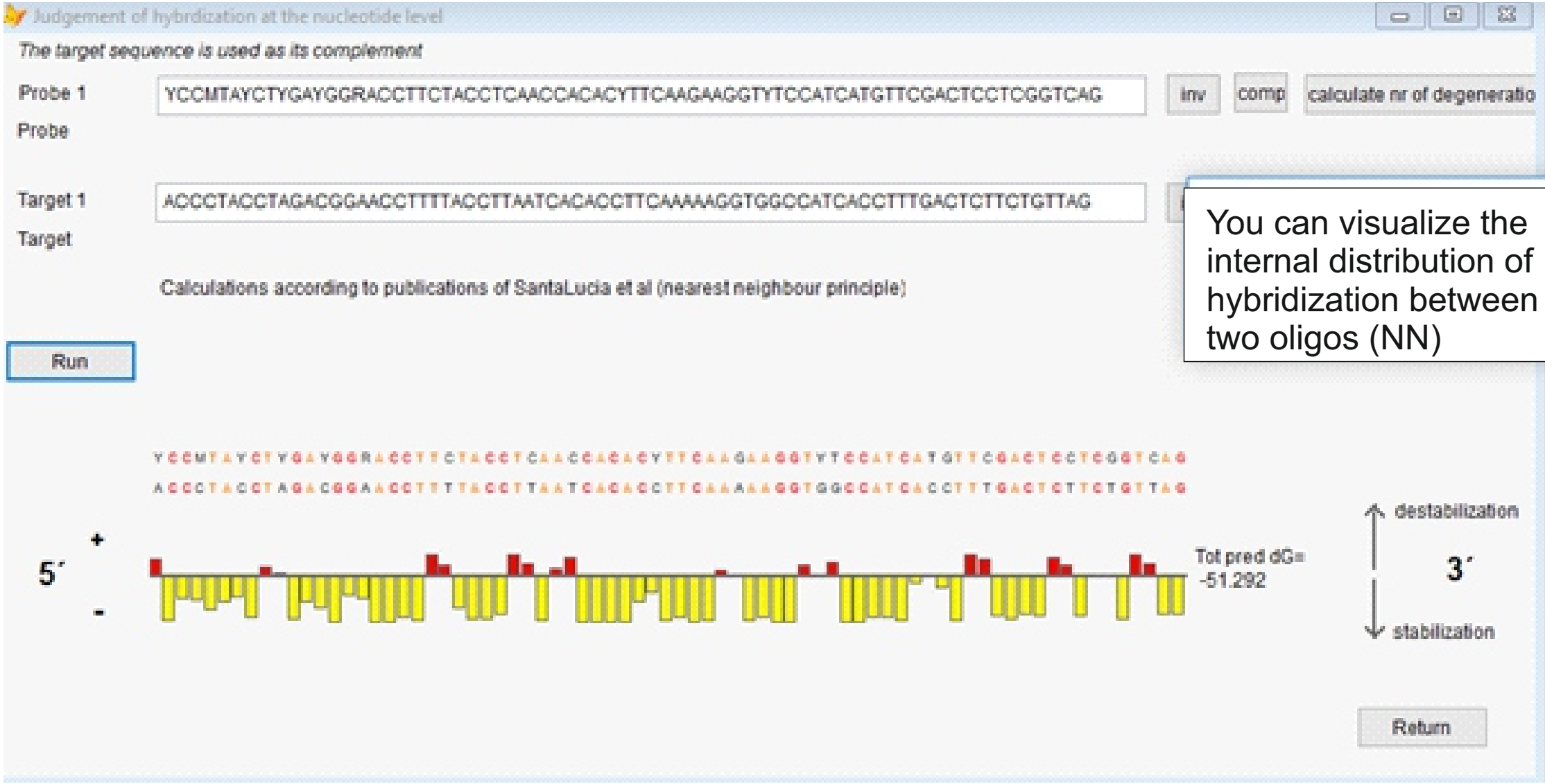
Useful information if you want fo construct primers and probes

# Test of a chosen primer/probe against haplotypes from a segment of the alignment

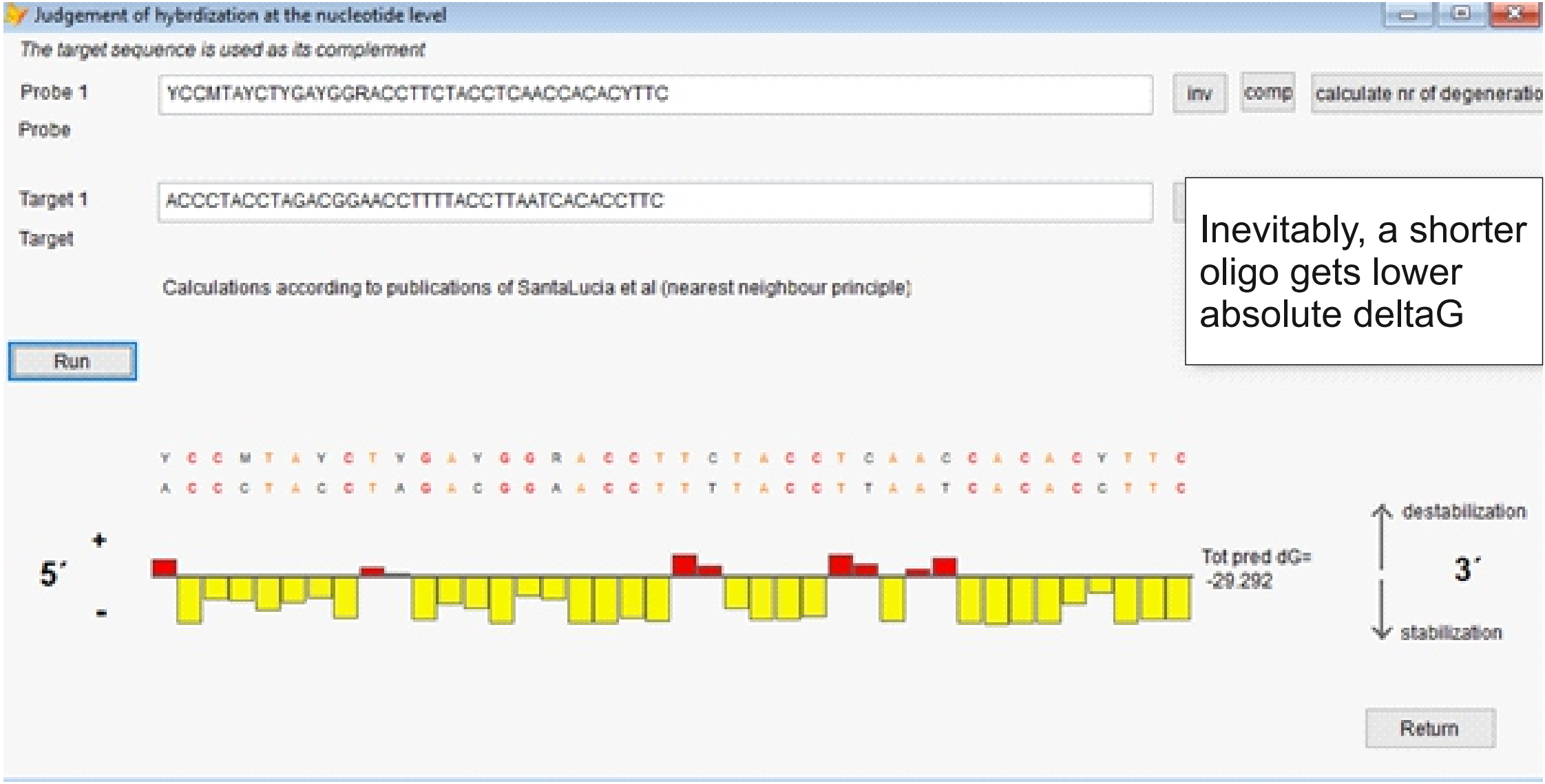


# Fine tuning of primer and probe properties

Inosine and degenerations can be added ad lib.

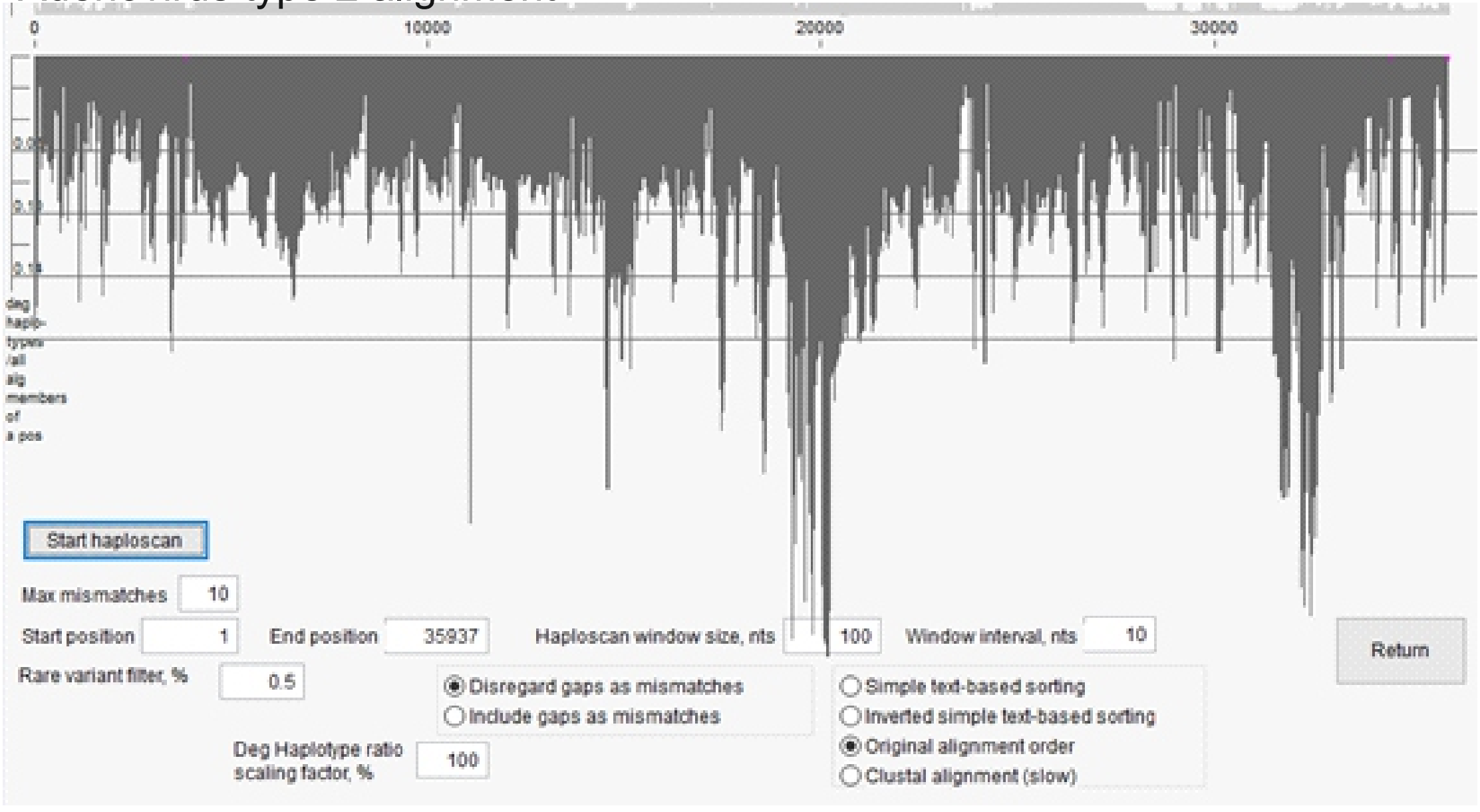


# Fine tuning of primer and probe properties



# Another way of looking at variability: Haploscanning

Adenovirus type 2 alignment



# Task 3: Implement PCR

JCM
Journals.ASM.org

## Variation-Tolerant Capture and Multiplex Detection of Nucleic Acids: Application to Detection of Microbes

**Christina Öhrmalm,[a] Ronnie Eriksson,[a,b] Magnus Jobs,[a,c] Magnus Simonson,[b] Maria Strømme,[d] Kåre Bondeson,[a] Björn Herrmann,[e] Åsa Melhus,[e] and Jonas Blomberg[a]**

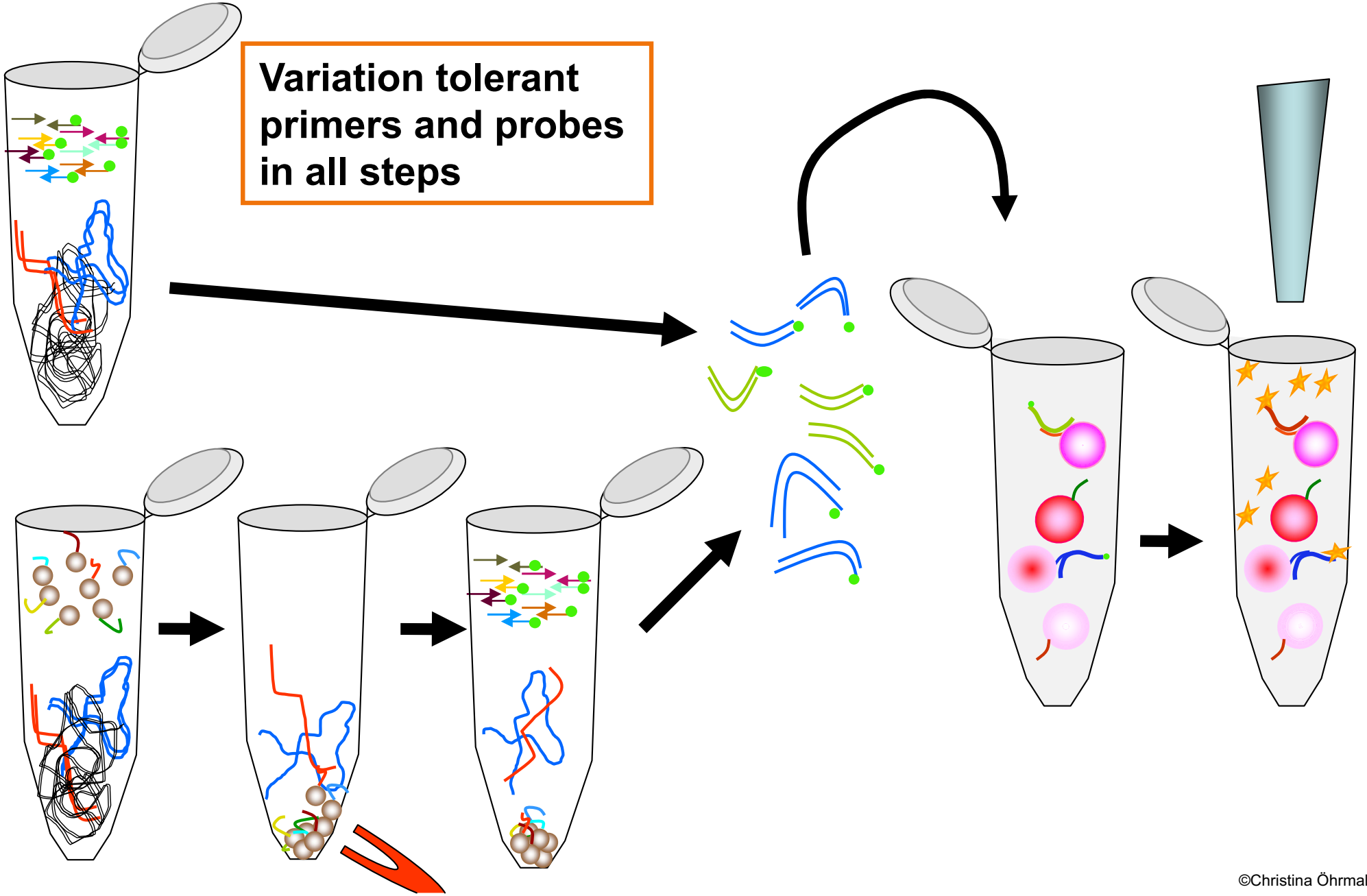
Sections of Clinical Virology[a] and Clinical Bacteriology,[e] Department of Medical Sciences, Uppsala University Hospital, Uppsala, Sweden; Science Department, National Food Agency, Uppsala, Sweden[b]; School of Health and Social Sciences, Högskolan Dalarna, Falun, Sweden[c]; and Nanotechnology and Functional Materials, The Ångström Laboratory, Uppsala University, Uppsala, Sweden[d]

**In contrast to ordinary PCRs, which have a limited multiplex capacity and often return false-negative results due to target varia-**

VOCMA

# VOCMA

## Variation Tolerant Capture Multiplex Analysis




©Christina Öhrmalm

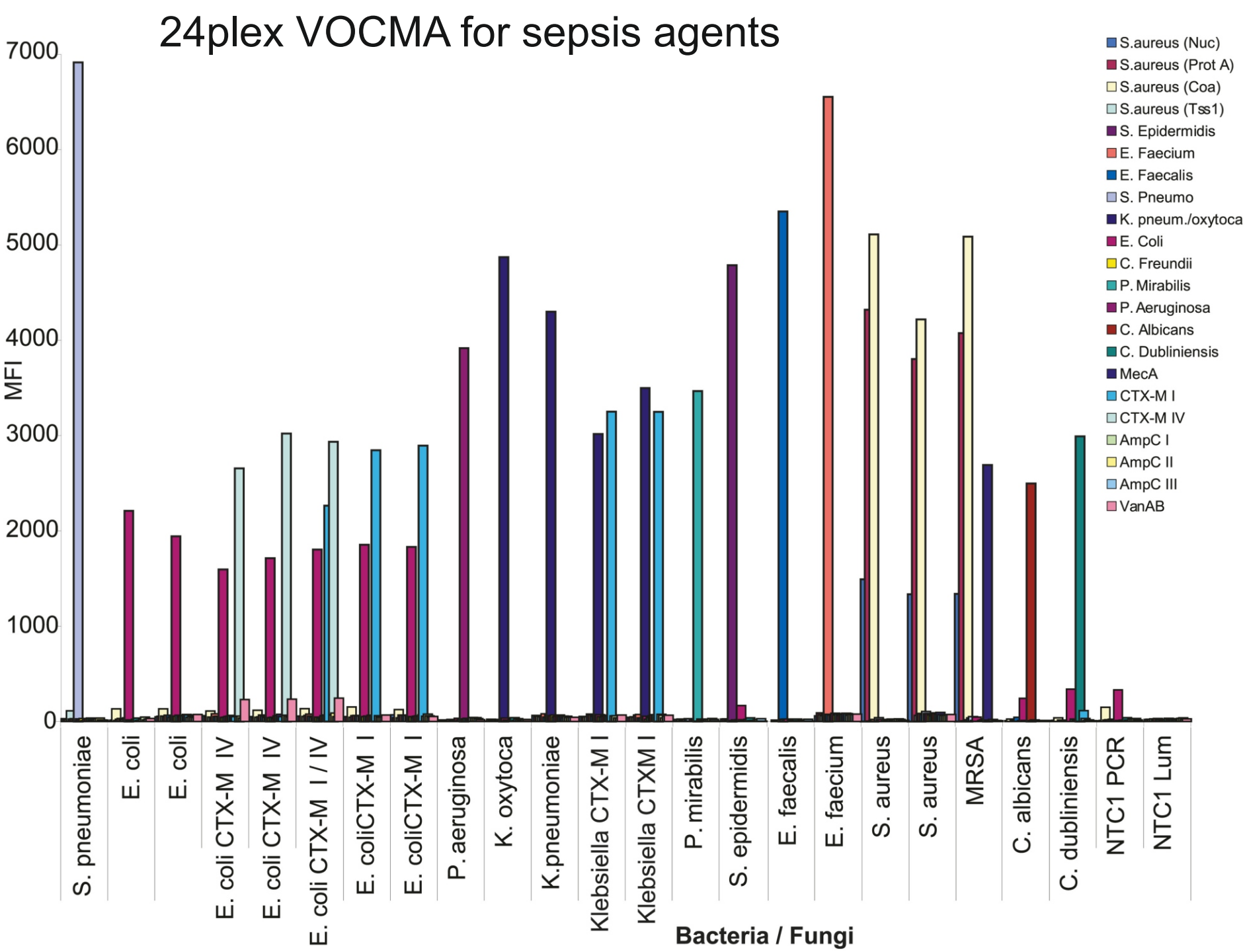
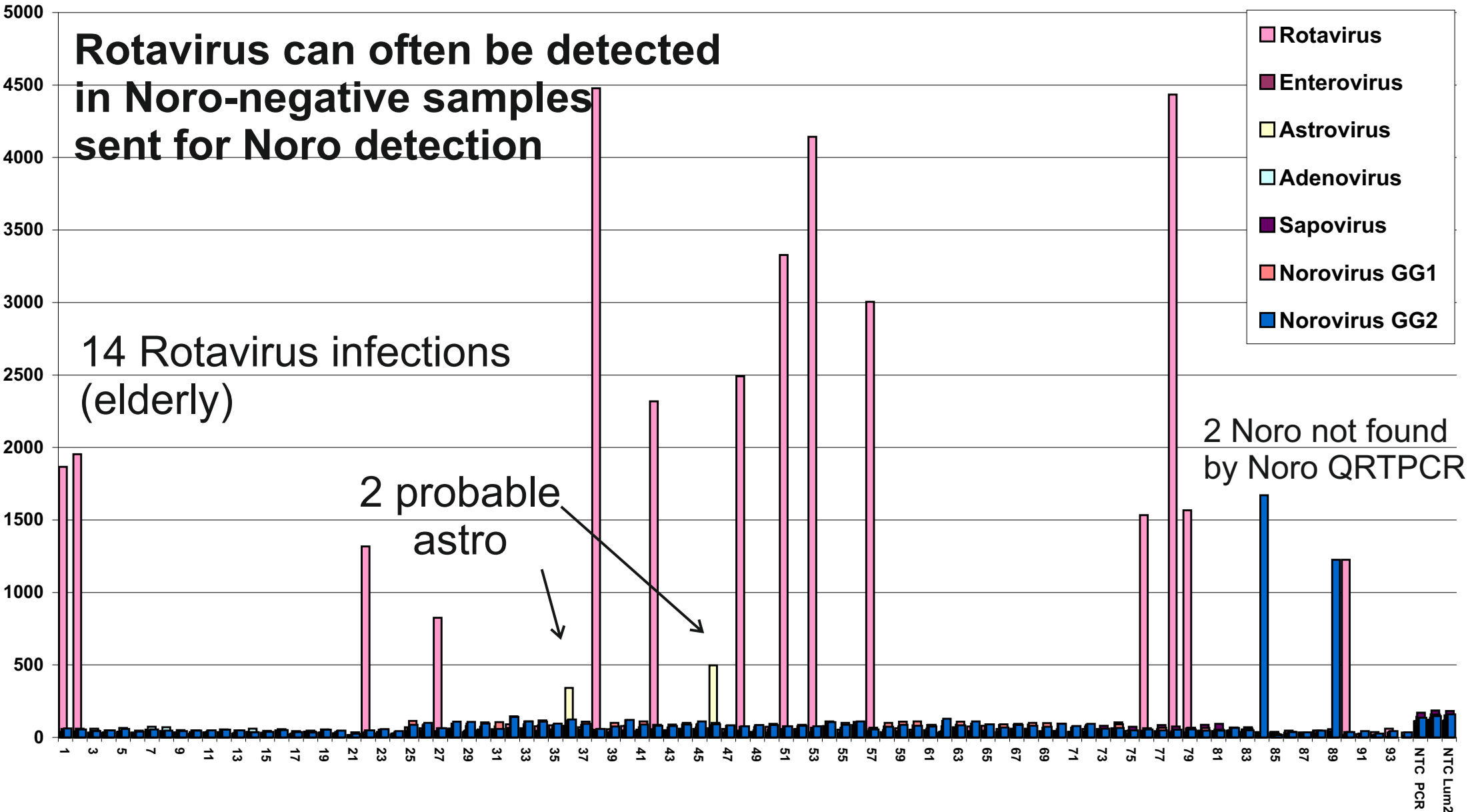



FIG 2 Specificity of sepsis VOCMA against cultured bacteria and fungi. A 22-plex VOCMA experiment on cultured strains of bacteria and fungi, with or without antibiotic resistance, from reference strains or strains from patient samples. The panel of 22 probes is indicated by different colors. The x-axis shows the cultures with different colors indicating different detection probes hybridizing. The y-axis shows the MFI.

# VOCMA

**94 gastroenteritis samples**

**(**tested negative in Norovirus GG1 and GG2 Q-PCR)

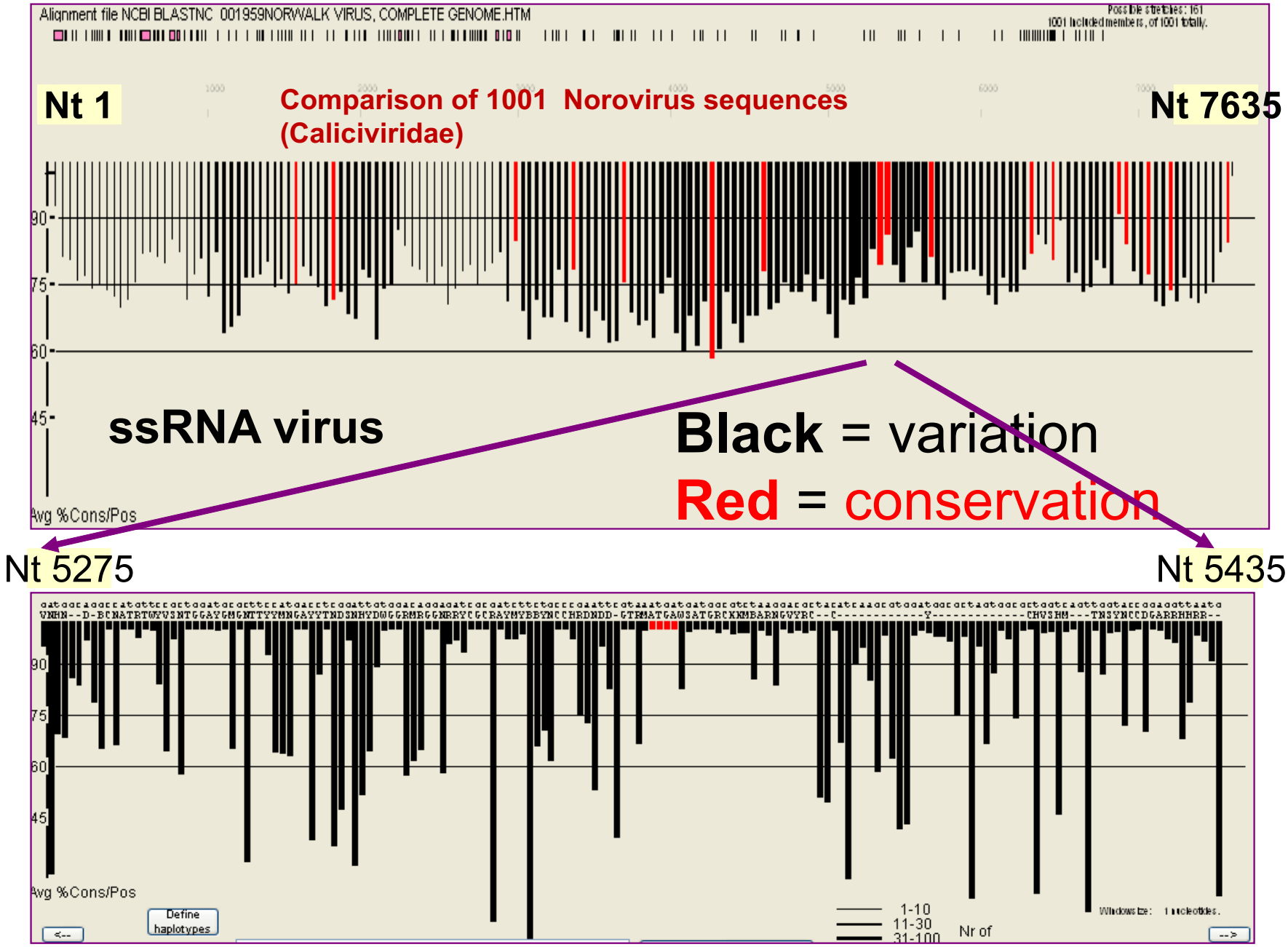


- In our hands: up to 24-plex

- requires opening of tubes
  after amplification,
  a contamination risk

We therefore developed a real-time PCR version of VOCMA

# Primer design: BlastN, Consort and Visual OMP




(NCBI Blastn and ConSort©, unpubl. J Blomberg)

© Christina Öhrmalm

# rtVOCMA

-variation tolerance with long primers and probes

-use of internal quencher in probe neccessary

-low specific primer concentration

-high generic primer concentration

-relatively high probe concentration

-in spite of check in Visual OMP,
we often have trouble with >5 plex

cDNA synthesis before PCR in case of RNA target

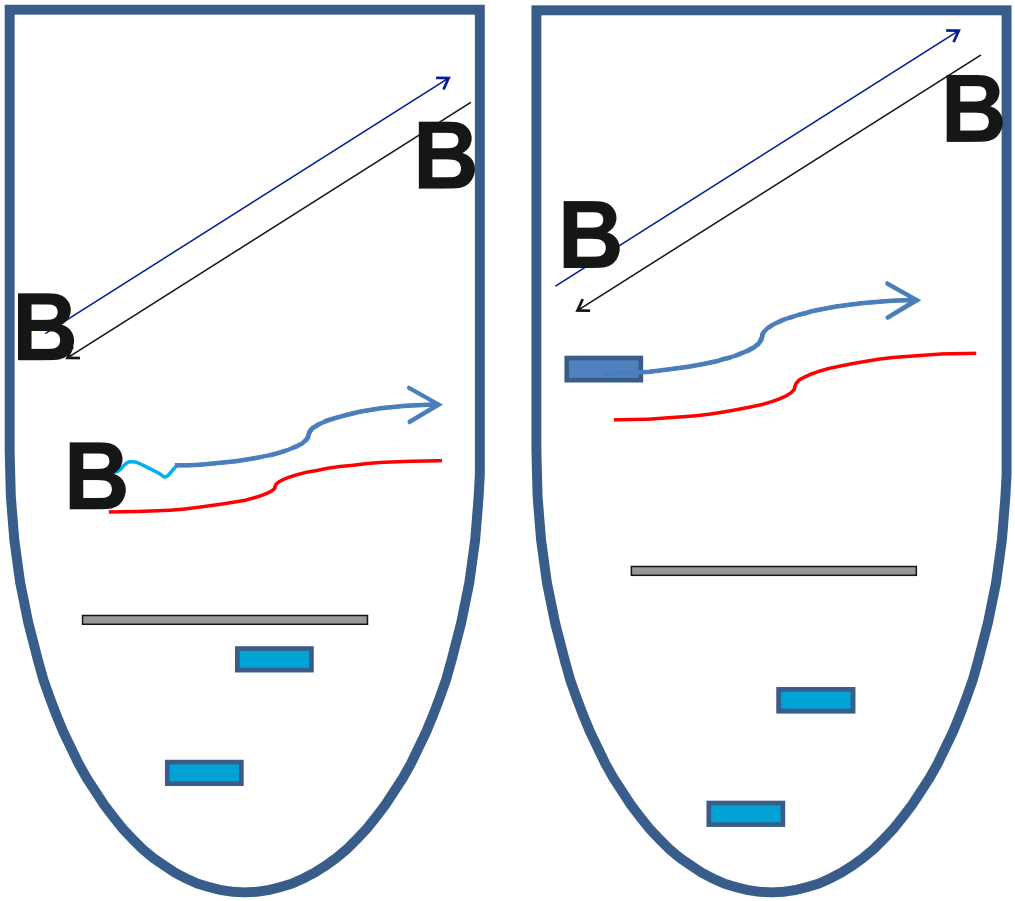
## Amplification according to single tube nesting principle*

### High temp amplification

Specific amplification
Low concentration of specific primers

### Lower temp amplification

Generic amplification
High concentration of generic primers




Generic primer

B

Biotinylated primer/probe

*Forsman et al , J Virol Methods 2003

# Amplification conditions

- Low concentration of specific primers:40nM
- Concentration of generic primers and probes: 300nM, 100nM

45°C 10 min

95°C 10 min

95°C 15 sec

61°C 45 sec

} 25 cycles, specific amplification by specific primers

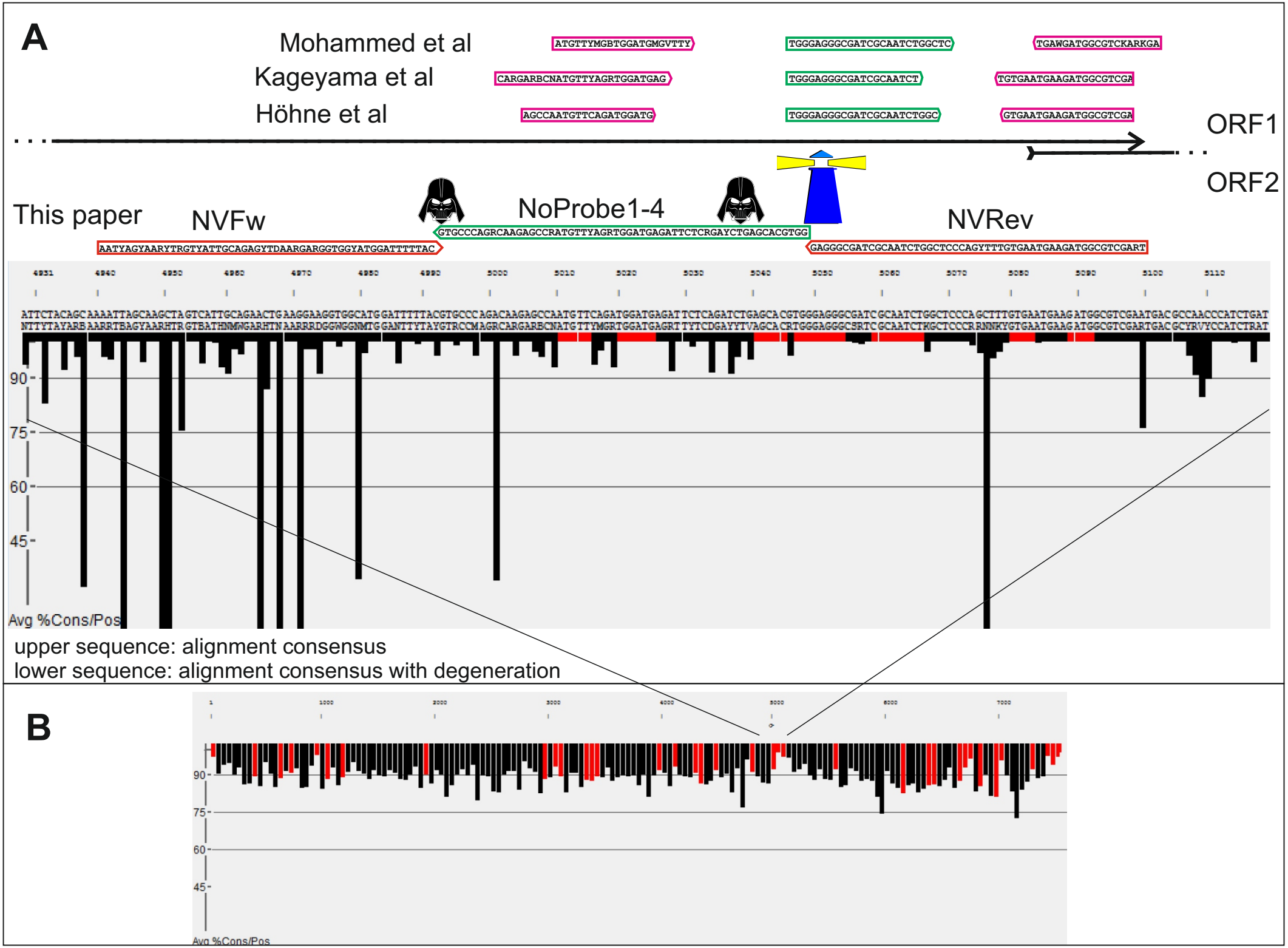
95°C 15 sec

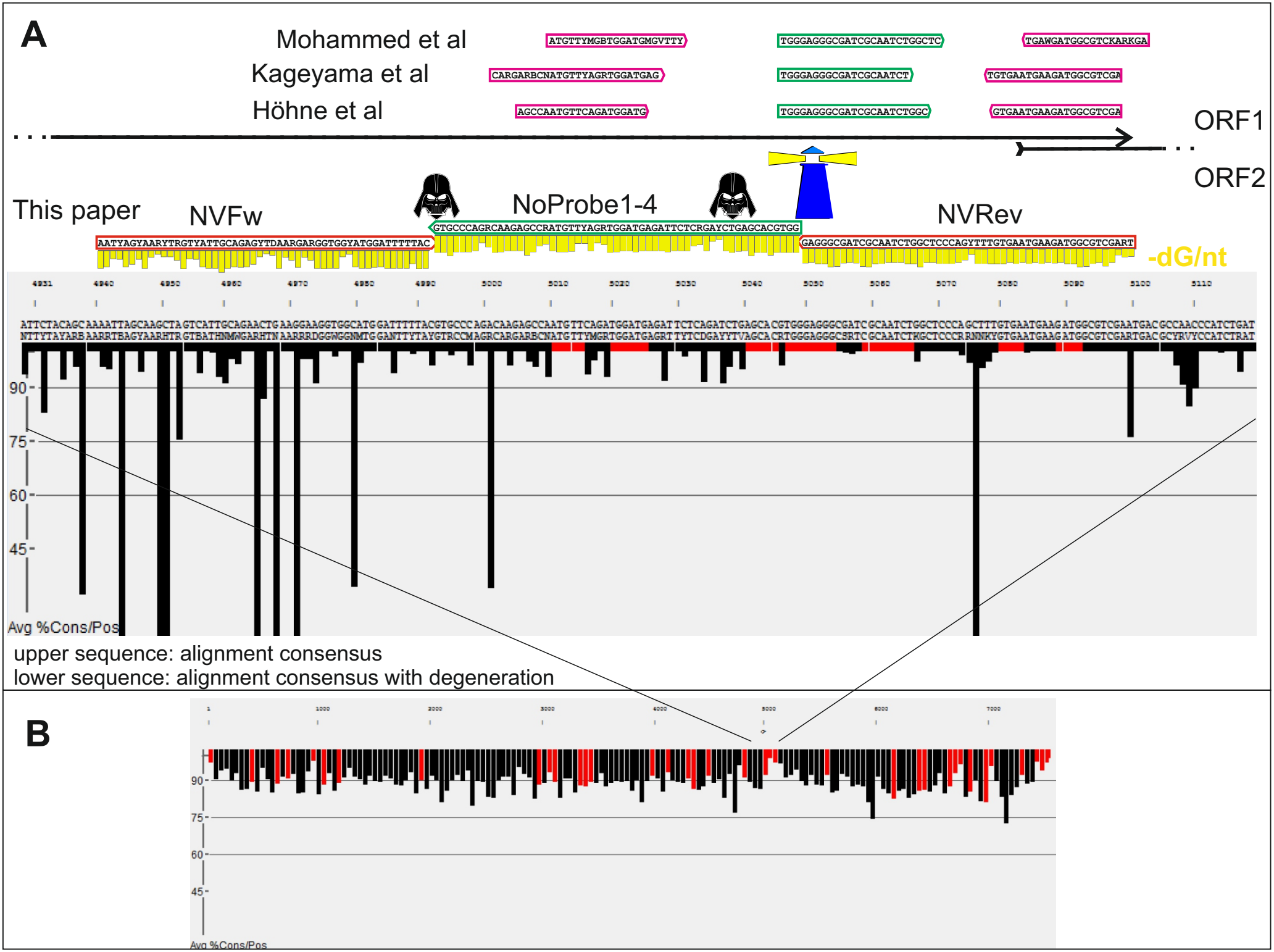
52°C 45 sec

} 40 cycles, amplification by generic primers

# Norovirus primers and probe, rtVOCMA

| Primers/Probes | Nucleotide sequences (5' - 3') |
|---|---|
| Generic first | TTGGATAAGTGGGATA |
| Generic second | AAGATATCGTAAGGAT |
| NVFw | AAGATATCGTAAGGATAATYAGYAARYTRGTYATTGCAGAGYTDAARGARGGTGGYATGGATTTTTAC |
| NVRev | TTGGATAAGTGGGATAAYTCGACGCCATCTTCATTCACAAARCTGGGAGCCAGATTGCGATCGCCCTC |
| Nvprobe1 | FAM/CCACGTGCTCAGRTCYGAGA/iBHQ1/ATCTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/3Spacer C3 |
| Nvprobe2 | FAM/CCACGTGCTCAGRTCYGAGAA[T(BHQ-1)]CTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/spacer C3 |
| NV probe3 | FAM/CCACGTGCT/zen/CAGRTCYGAGAATCTCATCCAYCTRAACATYGGCTCTTGYCTGGGCAC/3Iowa black RQ-Sp |
| Nvprobe4 | FAM/CCACGTGCT/zen/CAGRTCYGAGAATCTCATCC/zen/AYCTRAACATYGGCTCTTGYCTGGGCAC/3Iowa black RQ-Sp |


A
Mohammed et al
Kageyama et al
Höhne et al
ATGTTYMGBTGGATGMGVTTY
TGGGAGGGCGATCGCAATCTGGCTC
TGAWGATGGCGTCKARKGA
CARGARBCNATGTTYAGRTGGATGAG
TGGGAGGGCGATCGCAATCT
TGTGAATGAAGATGGCGTCGA
AGCCAATGTTCAGATGGATG
TGGGAGGGCGATCGCAATCTGGC
GTGAATGAAGATGGCGTCGA
ORF1
ORF2
This paper
NVFw
NoProbe1-4
NVRev
AATYAGYAARYTRGTYATTGCAGAGYTDAARGARGGTGGYATGGATTTTTAC
GTGCCCAGRCAAGAGCCRATGTTYAGRTGGATGAGATTCTCRGAYCTGAGCACGTGG
GAGGGCGATCGCAATCTGGCTCCCAGYTTTGTGAATGAAGATGGCGTCGART
Avg %Cons/Pos
upper sequence: alignment consensus
lower sequence: alignment consensus with degeneration
B
Avg %Cons/Pos


A
Mohammed et al
Kageyama et al
Höhne et al
ATGTTYMGBTGGATGMGVTTY
TGGGAGGGCGATCGCAATCTGGCTC
TGAWGATGGCGTCKARKGA
CARGARBCNATGTTYAGRTGGATGAG
TGGGAGGGCGATCGCAATCT
TGTGAATGAAGATGGCGTCGA
AGCCAATGTTCAGATGGATG
TGGGAGGGCGATCGCAATCTGGC
GTGAATGAAGATGGCGTCGA
ORF1
ORF2
This paper
NVFw
NoProbe1-4
NVRev
GTGCCCAGRCAAGAGCCRATGTTYAGRTGGATGAGATTCTCRGAYCTGAGCACGTGG
AATYAGYAARYTRGTYATTGCAGAGYTDAARGARGGTGGYATGGATTTTTAC
GAGGGCGATCGCAATCTGGCTCCCAGYTTTGTGAATGAAGATGGCGTCGART
-dG/nt
Avg %Cons/Pos
upper sequence: alignment consensus
lower sequence: alignment consensus with degeneration
B
Avg %Cons/Pos

# **Performance of the real-time version of VOCMA.**
## Noro GI and GII in one well, Rota and Sapo in another well

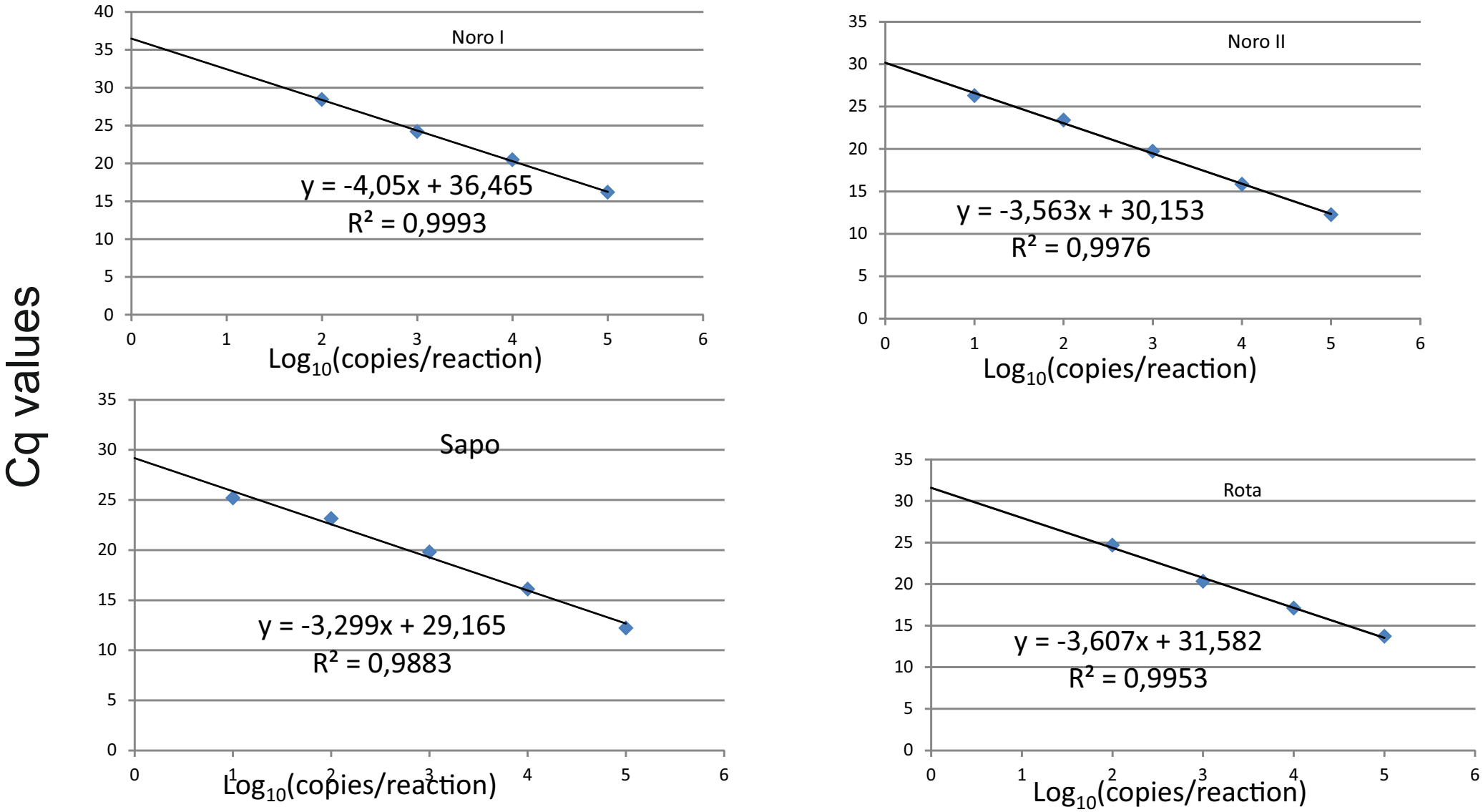



**Fig.1 Standard curves for the norovirus GII real-time PCR assay using 10-fold serial dilutions of the synthetic target.** The assays were performed in duplicate, and the average values were plotted. Only the Cq values from the second round of PCR were used.

| Virus | Detection limit (copies/reaction) | |
|---|---|---|
| | Singleplex | Duplex |
| Noro GGI | 10 | 100 |
| Noro GGII | <10 | <10 |
| Rota | 100 | 100 |
| Sapovirus | <10 | <10 |

## Conclusions

ConSort can handle large BLAST alignments,

-facilitate finding conservation,

-transform alignments into haplotypes,

-aid in design of degenerated primers and probes <100 nt.

rtVOCMA can be used for broadly targeted PCRs

# Future developments

Distributable program package (?)

Overcoming changes in BLASTN output format (NCBI)

Improved primer/probe fit search in alignments

Improved graphics

**Collaborators and investors who want to share the troubles and fortunes ahead are welcome!**

# Collaborators

Christina Öhrmalm*

Hongyan Xia*

Jacob Ottoson

*thanks for some illustrations!