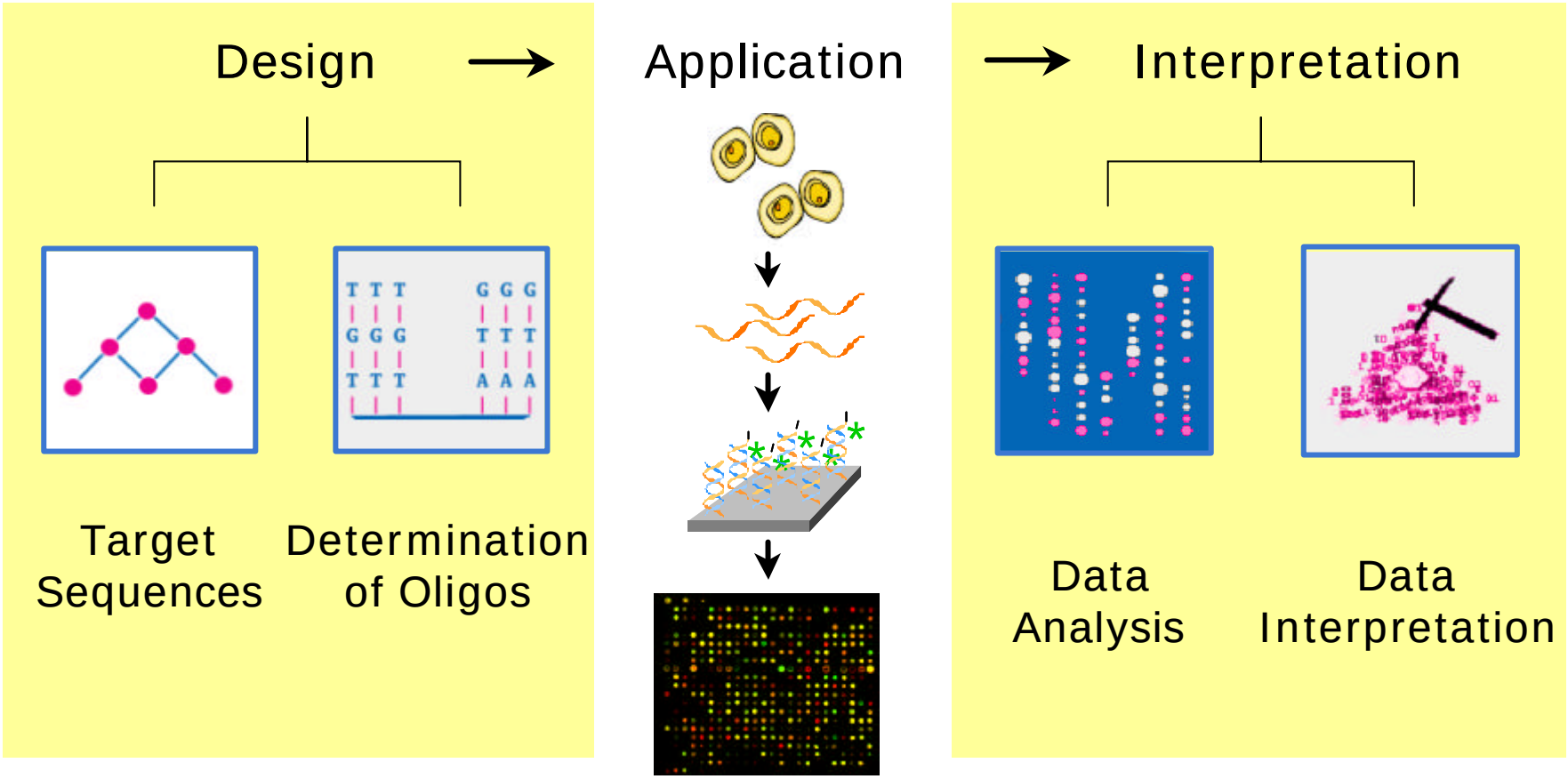
CENTRUM FÜR ANGEWANDTE GENSENSORIK

BIBIS ttz Bremerhaven

Bioinformatics Tools for Choosing the Right Capture Sequences

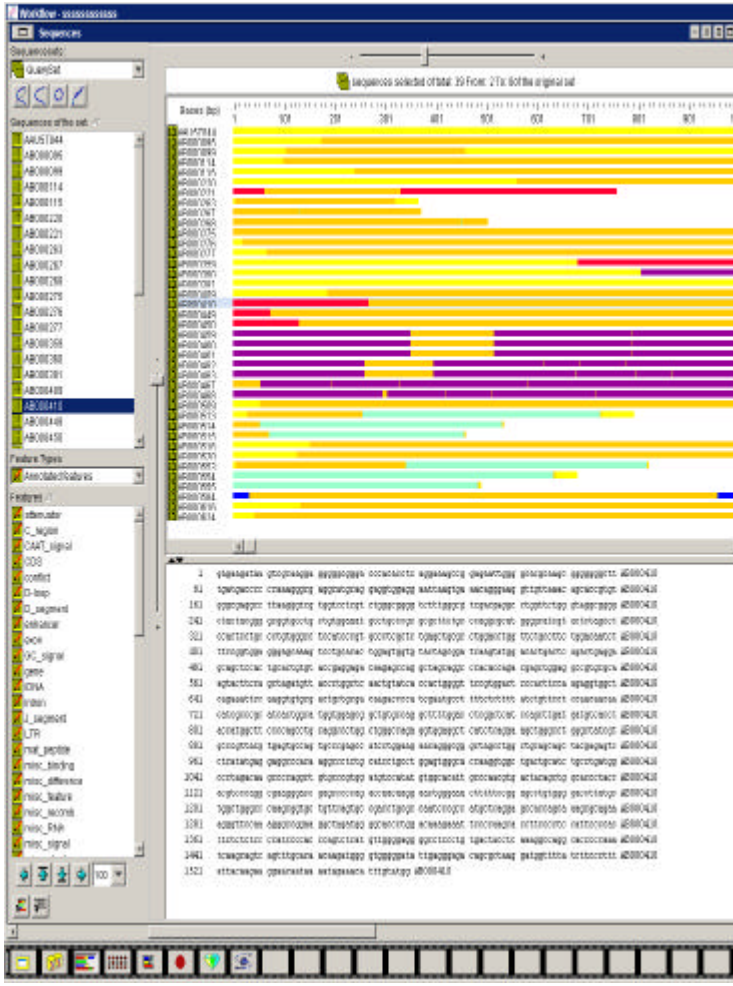
Uta Bohnebeck (TTZ-BIBIS)
Manfred Nölte (ZAIT)
Gerald Volkmann (iSenseIt)
Thomas Waschulzik (iSenseIt)
Tom Wetjen (TZI)

Workflow


Design
Application
Interpretation
T T T
G G G
T T T
G G G
T T T
A A A
Target Sequences
Determination of Oligos
Data Analysis
Data Interpretation

Determination of Target Sequences

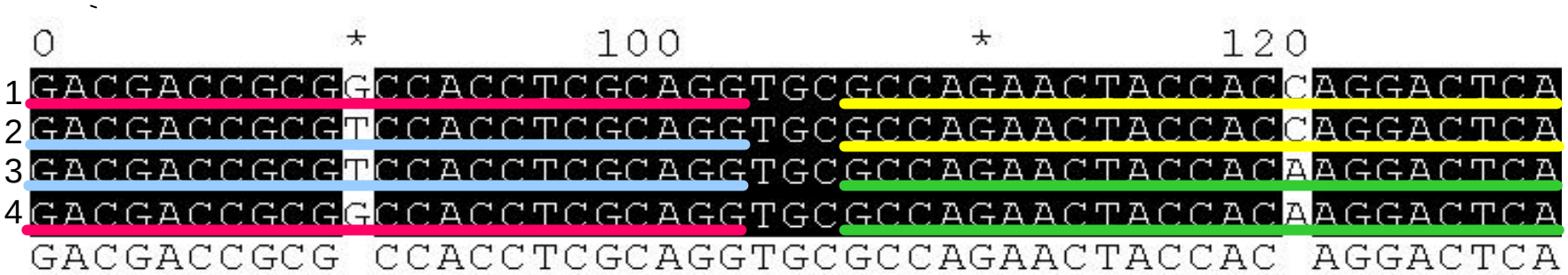
- Taxonomy-based selection of target sequences
- Relative arrangements of the selected sequences
- Alignments
- Text views
- Combined search queries

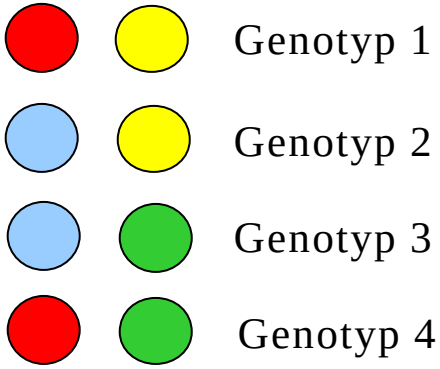


Computation of the Capture Oligos

- Consideration of the following criteria:
  - High sensitivity and specificity
  - Range of melting temperature
    - Length of capture oligos
    - GC-content
    - Salt concentration of hybridization buffer
  - Secondary structure of the capture oligos and target sequences
- Methods
  - Intelligent combinatorics
  - Optimization by a greedy algorithm
  - Genome-wide specificity test based on „thermodynamic alignments“
- Results:
  - Hierarchical oligo libraries

Example for Intelligent Combinatorics





	Sample A	Sample B	Sample C	Sample D
	+	-	-	+
	+	+	-	-
	-	-	+	+
	-	+	+	-
Genotyp	1	2	3	4

Optimized Specificity Test

- Example: Mismatch-Distance 3

5`ACTGGCGTGGAATTGTGACC 3` Target sequence

3`TGACCGCACCTTAACACTGG 5` Capture oligo

5`ACTGACGTGCAATTGTAACC 3` non target sequence with mismatch-distance 3 = 85% Similarity

- „Thermodynamic alignment“:
  - Consideration of the relative position of mismatches
  - Consideration of mismatch type

Comparison with Other Approaches

ACTGGCGTGGAATTGTGACC

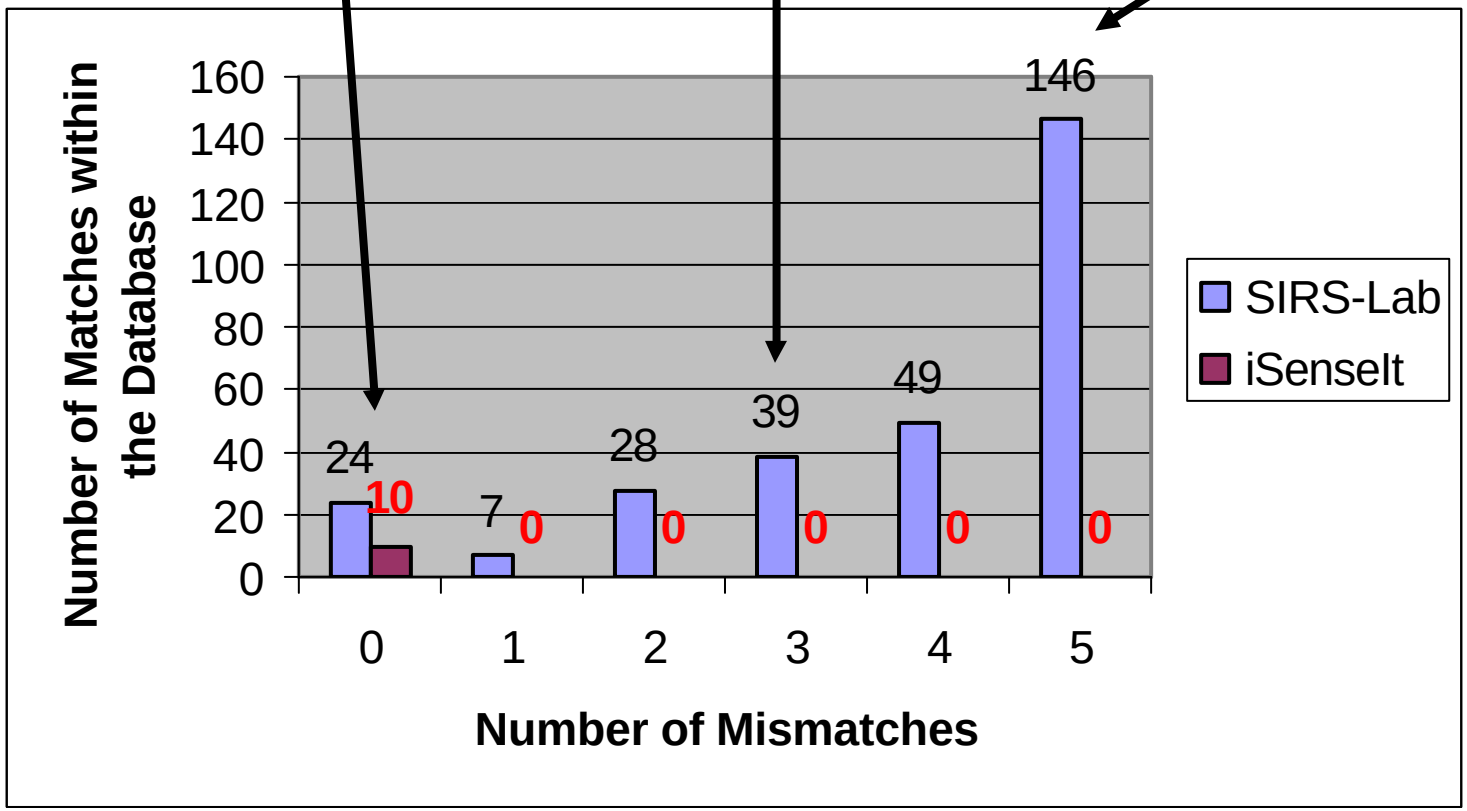
TGACCGCACCTTAACACTGG

ACTGACGTGCAATTGTAACC

TGACCGCACCTTAACACTGG

ACTGGAGTCCAATTAAGACC

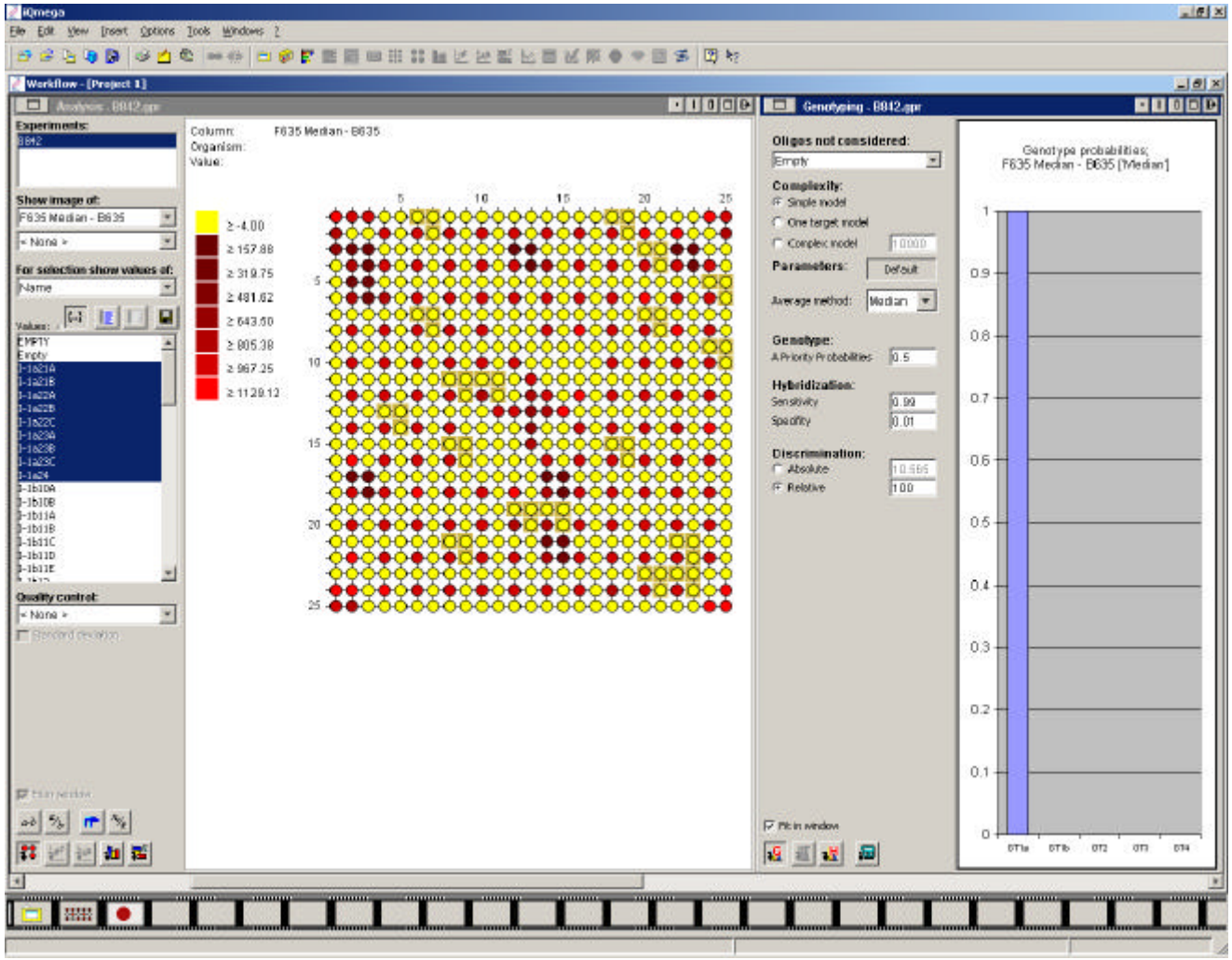
TGACCGCACCTTAACACTGG



Automated DNA-Chip Interpretation

- Interpretation of complex hybridization patterns
  - Knowledge-based approach
  - Consideration of uncertainty
- Consideration of all available information
  - Chip design
  - Redundance
  - Cross hybridizations

Genotyping of HCV



Intelligent Data Interpretation

- Question: Which HCV genotype is in the sample?
- Problems:
  - Unspecific hybridizations
  - False positive and false negative signals