Cancer research

R package to analyze genomic alterations and tumor pathways based on array data from single nucleotide polymorphism (SNP) and comparative genomic hybridization (CGH) experiments

Eberhard Korsching
University Hospital of Münster
Gerhard-Domagk-Institute of Pathology

Our research focus

Cancer development / progression
(e.g. Breast, Ewing's Sarcoma, Osteosarcoma)
• Prognostic / therapeutic factors
• Analysis of the regulatory system on the level of DNA, RNA and proteins based on
• Comprehensive sample archive
• Lab techniques like: TMA, Affymetrix 4C, TaqMan, Cell culture

Development of analysis solutions on this research background
Core platform : S-Plus – Fortran, now establishing R – Fortran

Design

nice to have:
• a data browser like in S-Plus for the workspace content
• more concern on big data sets > 600 MB
• R to Fortran translator for time critical calculations - or similar

From S-Plus to R – Reasons:
• Community
• Technical shortcomings – e.g. S-Plus has memory leaks
Task – migrating from S-Plus to R:
Primarily the graphics routines have to be adapted

data sets:
- parameters / annotations
- array data

one-to-many

SNP.envelope, multi
gene.dosage, a
SNP.cn.envelope
gene.dosage.indi
adapt.exprSet.toSNP
cutoff.peaks
SNP.envelope
plot.bar, point, segment
plot.chromosome, outline

Biology – SNP Copy Number Analysis

Genomic sequence
Control: T A A A C G G
Sample: T A A A C G G

Control (~100 samples)
because of signal fluctuations, and fluctuations of the base type in the population

Intensity

Intensity

A T G C
or
or

A T G C

Chromosome 4
raw SNP copy number, Mapping 10K Affymetrix A431 cell line
**Visualization**

SNP copy number across genome

MCA-MB-468 cell line, Mapping 10K Affymetrix, smoothing window: 40

The colored area indicates genetic alterations - Gains: green, losses: red

**Take home message**

- S-Plus to R is an easy task
- SNPs are capable to replace the CGH technique
- Old CGH data can be integrated

Improvements in the Analysis Strategy Make Single Nucleotide Polymorphism Analysis a Powerful Tool in the Detection and Characterization of Amplified Chromosomal Regions in Human Tumors

Eberhard Korsching a Konstantin Agelopoulos a Hartmut Schmidt a Inka Buchroth a Georg Gosheger b Pia Wülfing c Werner Böcker a Burkhard Brandt a Horst Buerger a

Institutes of a Pathology and b Orthopedics and c Department of Gynecology, University of Münster, Münster, and d Institute of Tumor Biology, University of Hamburg, Hamburg, Germany

Pathobiology 2006;73: (DOI:10.1159/000093088)

Cooperation with: Walter Nadler
Complex Systems Research Group, John von Neumann Institute for Computing, Research Centre Jülich, Germany & Computational Nano- and Biophysics Group, Department of Physics, Michigan Technological University, USA

**Biology – CGH vs. SNP Analysis**

Comparative Genomic Hybridisation

- Test DNA Fluorochrome 1
- Reference DNA Fluorochrome 2

Hybridize

Metaphase Chromosome Spread

Ratio of the intensities of two fluorochromes along the target chromosome indicate regions of genetic gain and loss
Results I

CGH smoothing window 50
SNP smoothing window 20
SNP smoothing window 1

Results II

100% tumour DNA
80% tumour DNA
60% tumour DNA
40% tumour DNA
20% tumour DNA

Chr. 18 SNPs: 346
Chr. 18 SNPs: 346
Chr. 18 SNPs: 346
Chr. 18 SNPs: 346
Chr. 18 SNPs: 346