CAIRO User Story

User	Evert van den Broek
CTMM Project(s)	DeCoDe
Role in project(s)	Researcher
Goal	To validate MSI as a prognostic biomarker for disease progression in metastatic colorectal cancer
Benefit	Combine experimental and clinical data to perform analyses (survival- and frequency plots), and to make data publicly available

Steps to answer research question (current situation)

1. Acquire clinical data (Excel file)
2. Tumor classification (TNM staging, MSI high/low) by histopathology
3. Acquire genomics data, for CNV (aCGH)
4. Run CGHcall pipeline
5. Map segmentdata back to probes, using an R-script (reads files with probenames and segmentdata and writes a file with segmentdata per probe, per sample)
6. Create frequency plots in R (gains/losses per chromosome)
7. Create survival analysis plots on combined clinical and biomarker data

Desired situation

1. Clinical data available, together with biomarker data
2. Uploading genomics data into the integrative WP5 tool
3. Run CGHcall pipeline from within this tool
4. Run the pipeline (R-package) for backmapping segementdata to probes, also from within this tool.
5. Combine clinical and biomarker data to perform survival analyses and frequency plots

Dataset

A large study on genomic profiling and survival times, in colorectal cancer patients. This was done by classification of MSI (low or high MicroSatellite Instability),) and acquiring genomic data, where CNV is measured using aCGH.
Various other clinical variables and biomarkers were measured, such as treatment and surgery information, tumor classification, survival time in months.

Cohort information

CAIRO 1 study: 820 colorectal cancer patients in 74 Dutch hospitals, in 24 months.

Clinical data

The clinical data is available in Excel, but is also being stored into the TraIT OpenClinica instance.

Biobanking

Tumor tissue samples.

Experiment data

Tumor samples were classified for low or high MSI, and genomics data was acquired CNV calculation. This was done using aCGH segmentation and calling.

Research Question

What is the prognostic value of genetic profiles, low and high MSI, in CRC liver metastases and their corresponding primary tumors?

Example workflow files

Excel file containing clinical data

CAIROnr	gender	treatment arm	Date of randomisation	Age	Best response according to RECIST 1st line	PFS1	PFS1event	Overal survival	Cause of death	Site primary tumor	Location metastases	MSI		
1	F	Sequential (A)	1.1.1900	58	PR	60								
2	F	Combination (B)	1.1.1900	73	PR	120								
3	M	Combination (B)	1.1.1900	60	early death toxicity	548								

Input R-script: Call data per probe, tumor versus normal tissue (tab-delimited)

Probe name	Chromosome	Start	Cairo.12.T.vs.N	Cairo.13.T.vs.N	Cairo.14.T.vs.N	Cairo.15.T.vs.N	Cairo.16.T.vs.N	etc						
A_16_P00000027	1	784458	-1	0	0	-1	-1							
A_16_P00000036	1	799631	-1	0	0	-1	-1							
A_16_P00000037	1	802868	-1	0	0	-1	-1							
etc														

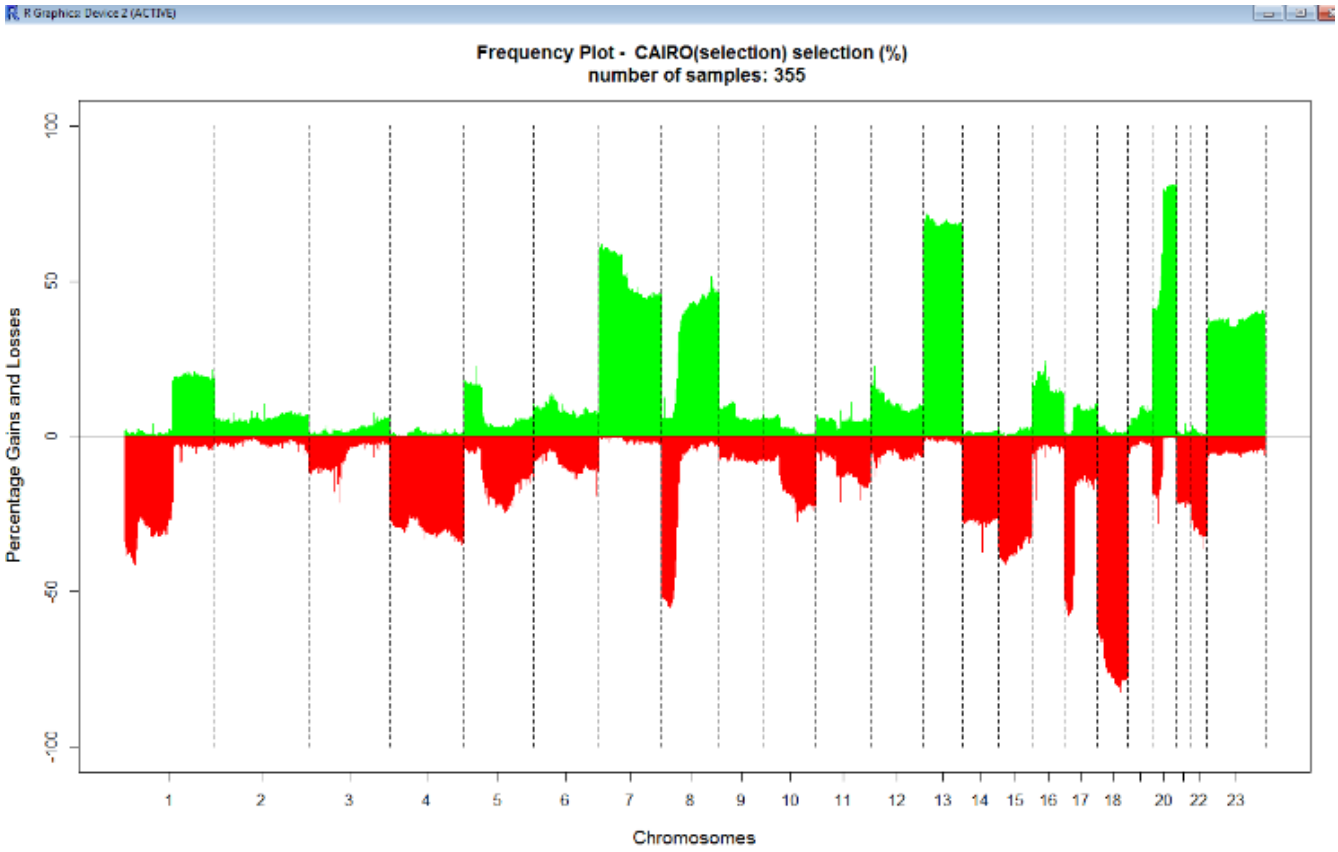
Input R-script: values per segment

Sample	values	start	end	chrs	nclone					
Cairo.2.T.vs.N	-0.211	1	10757	1	10757					
Cairo.2.T.vs.N	0.177	10758	10762	1	5					
Cairo.2.T.vs.N	-0.212	10763	13275	1	2513					
Cairo.2.T.vs.N	0.205	13276	13281	1	6					
Cairo.2.T.vs.N	-0.191	13282	13714	1	433					
Cairo.2.T.vs.N	-0.003	13715	28028	2	14314					
etc										

Output R-script: segmented CNV values, mapped back on probes

Probe name	Chromosome	Start	Cairo.2.T.vs.N	Cairo.3.T.vs.N	Cairo.4.T.vs.N	Cairo.5.T.vs.N	Cairo.6.T.vs.N	etc						
A_16_P00000027	1	784458	-0.211	0.003	0.041									
A_16_P00000036	1	799631	-0.211	0.003	0.041									
A_16_P00000037	1	802868	-0.211	0.003	0.041									
etc														

Output R: Frequency plot, gains losses per chromosome


R Graphics: Device 2 (ACTIVE)
Frequency Plot - CAIRO(selection) selection (%)
number of samples: 355
Percentage Gains and Losses
100
50
0
-50
-100
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
20
22
23
Chromosomes