Quality control for Affymetrix GeneChips®: What MM's are good for

Different aspects of quality control

What Affymetrix and Bioconductor offer

Evaluation & conclusion

Oliver Hartmann, Birgit Samans

Institute of Medical Biometry and Epidemiology,
Philipps University Marburg, Germany

Workshop on Statistical Analysis of Gene Expression Data,
11-14 July 2003, Wye College Conference Center, Kent, UK

Quality control: something we need to look at?

- there are standards for quality control
  - tests for chip quality
  - tests for tissue classification and purity
  - tests for mRNA quality
  - tests for hybridisation experiment

⇨ easy to detect poor quality in the lab

- life is not that easy:
  - quality control tests are often skipped
  - not sensitive enough
  - everything that can go wrong will go wrong

Two aspects of quality control: detecting poor hybridisation and outliers

1. reasons for poor hybridisations
   - mRNA degenerated
   - one or more experimental steps failed
   - poor chip quality, …

2. reasons for (biological) outliers
   - infiltration with non-tumour tissue
   - wrong label
   - contamination, …

✚ outlier detection for individual genes
   - due to dust, scratches, local effects, etc.
   - not discussed here

Quality control: a question of balance

- excluding too many reduces power
  - as sample size decreases
  - or increases costs if measurements are repeated
- including too many reduces power
  - introduction of additional noise
- defining the right quality measure not trivial
- choice of threshold might depend on
  - data set & sample size, experimental question, analysis method applied
- focus of this talk: hybridisation quality

Affymetrix-Guidelines for assessing sample and array quality: MAS 5.0

hybridisation quality

- probe array image inspection
- average background / noise values
- control spots
  - B2 oligos, poly-A controls, hybridization controls, internal control genes (GAPDH)
- percentage of present genes
  - test whether PMs are significantly bigger than MMs
- scaling & normalization factor

⇨ general guideline: values comparable

⇨ no methods for outlier detection

Quality control methods implemented in Bioconductor, R or standard software

- density plots of PM intensities
  - compare density distribution between chips
  - motivation: RMA rescales to average density
- image plot
- RNA degradation plots
  - to identify trends from 5‘ to 3‘ end
  - motivation: degradation starts at 5‘ end
- MvA plots for all pairs
- correlation/ scatter plot matrix between chips
- principal components analysis
- …

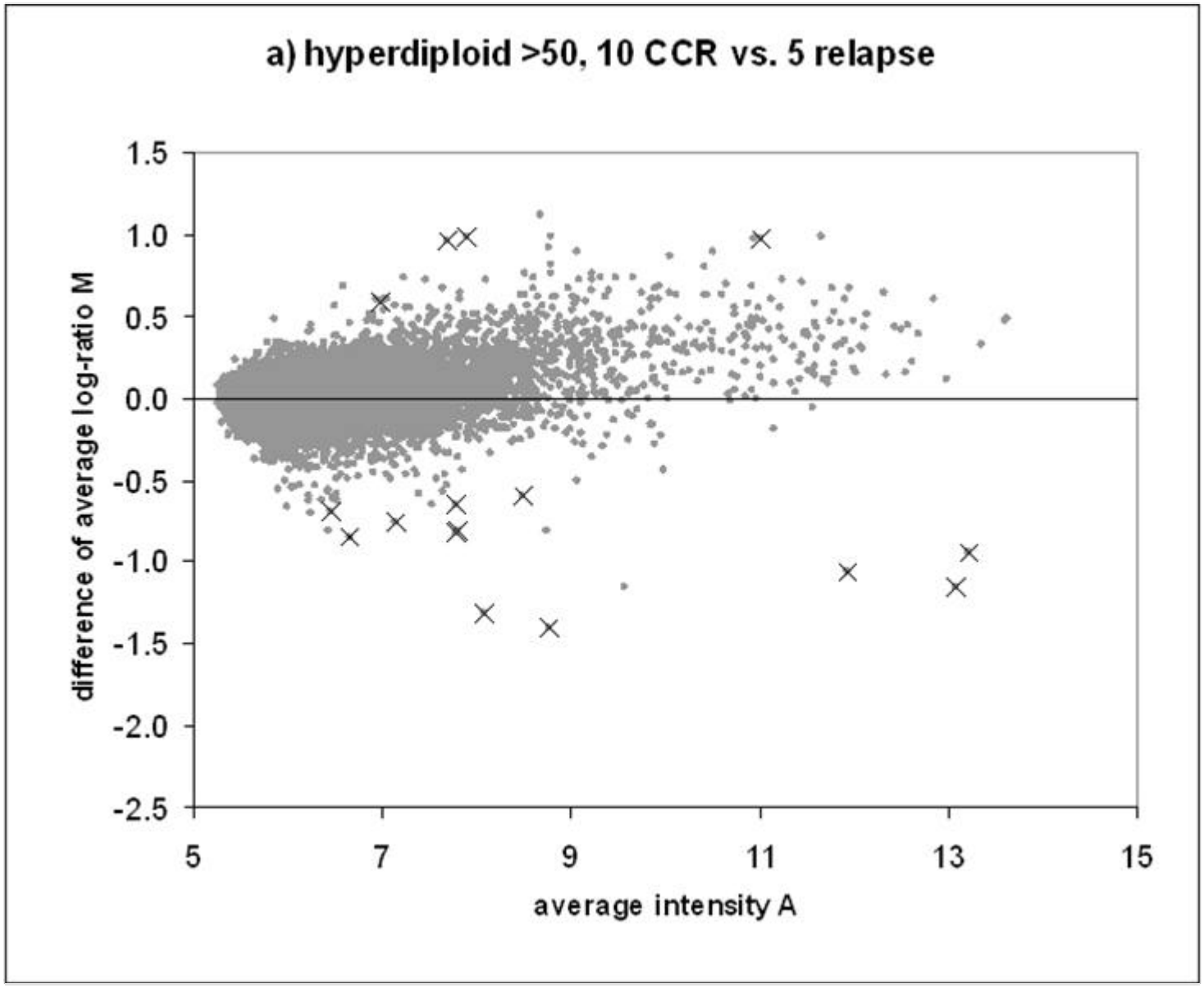
Example 1: reanalysis of a published data set

- two sample test, four subgroups
  - CCR vs. relapse, paediatric ALL subgroups
  - small sample size for relapse
  - RMA + fold change vs. MAS 5.0 + t-statistic
- app. 7% of hybridisations already excluded
  - 29 of 389 microarrays
- „high-quality gene expression data“
  - from HG_U95Av2

⇨ Is additional quality control necessary?

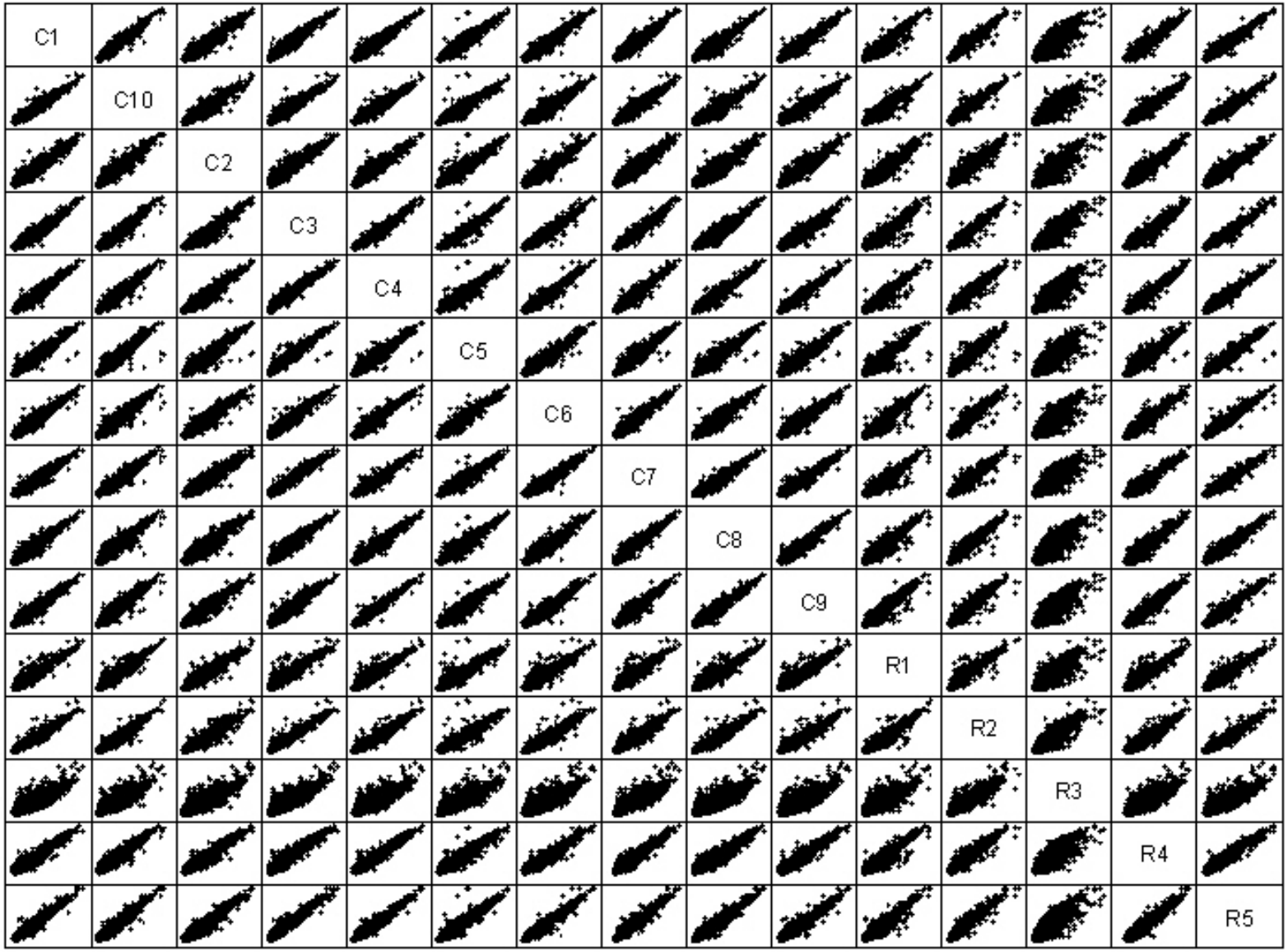
After RMA normalisation, resulting summary MvA plot does not behave as assumed

hyperdiploid > 50

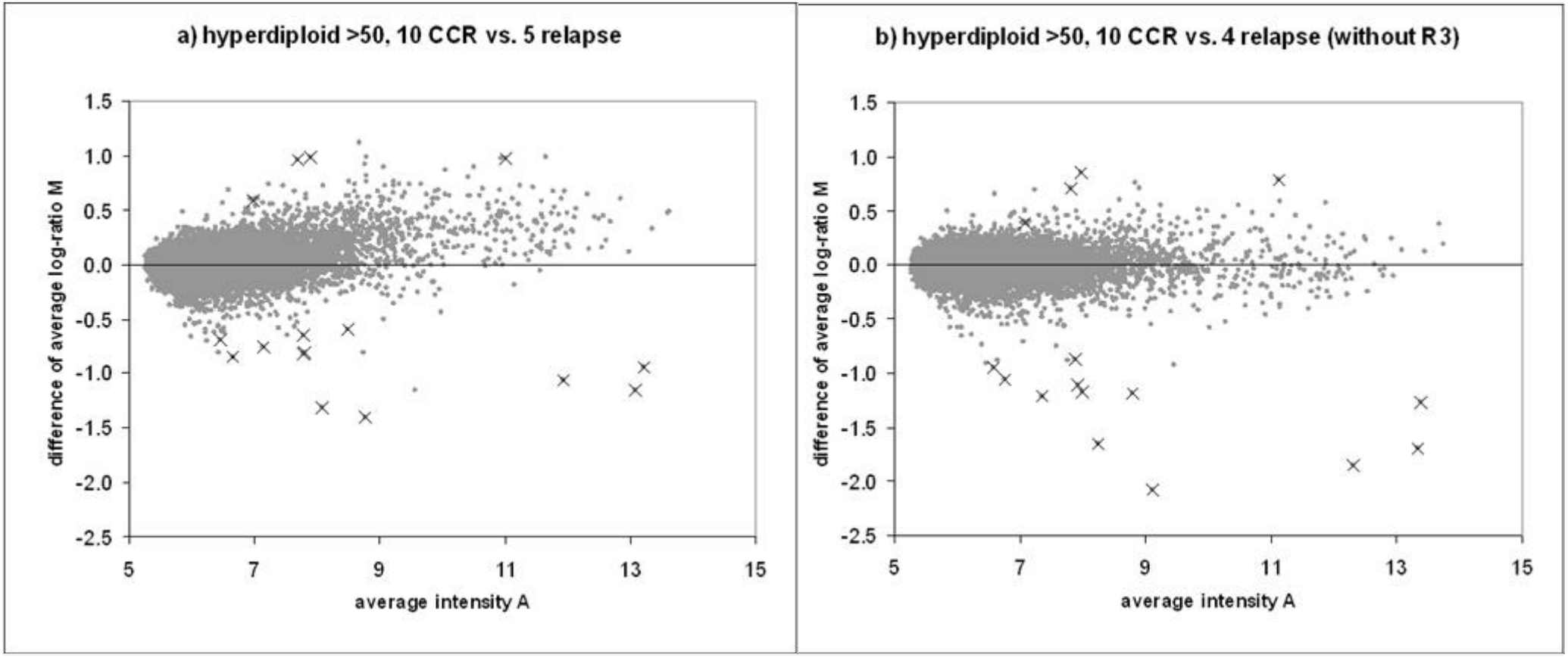


Scatter plot matrix & density plot detect outlier

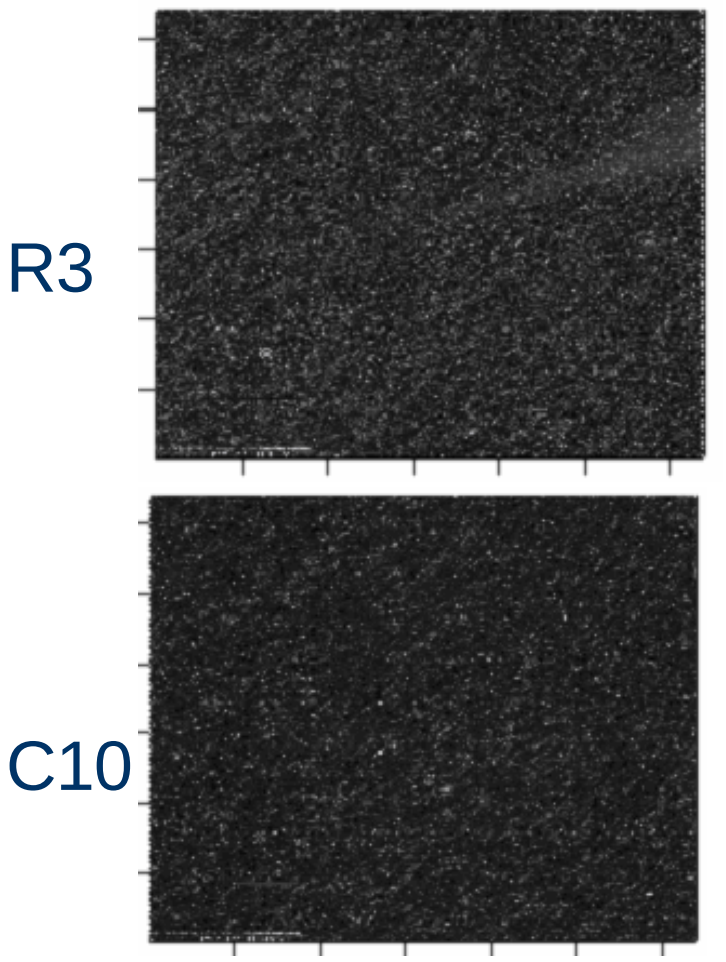
hyperdiploid > 50



Removing the outlier (R3) significantly improves results



Poor hybridisation quality or biological outlier?

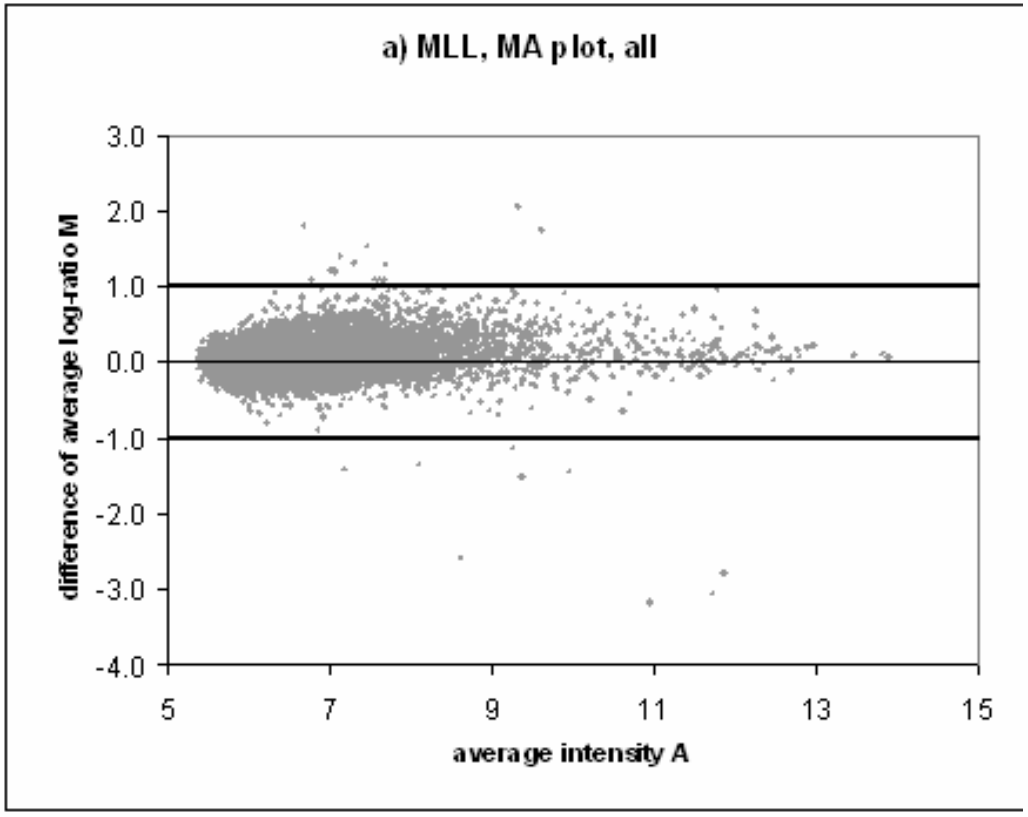


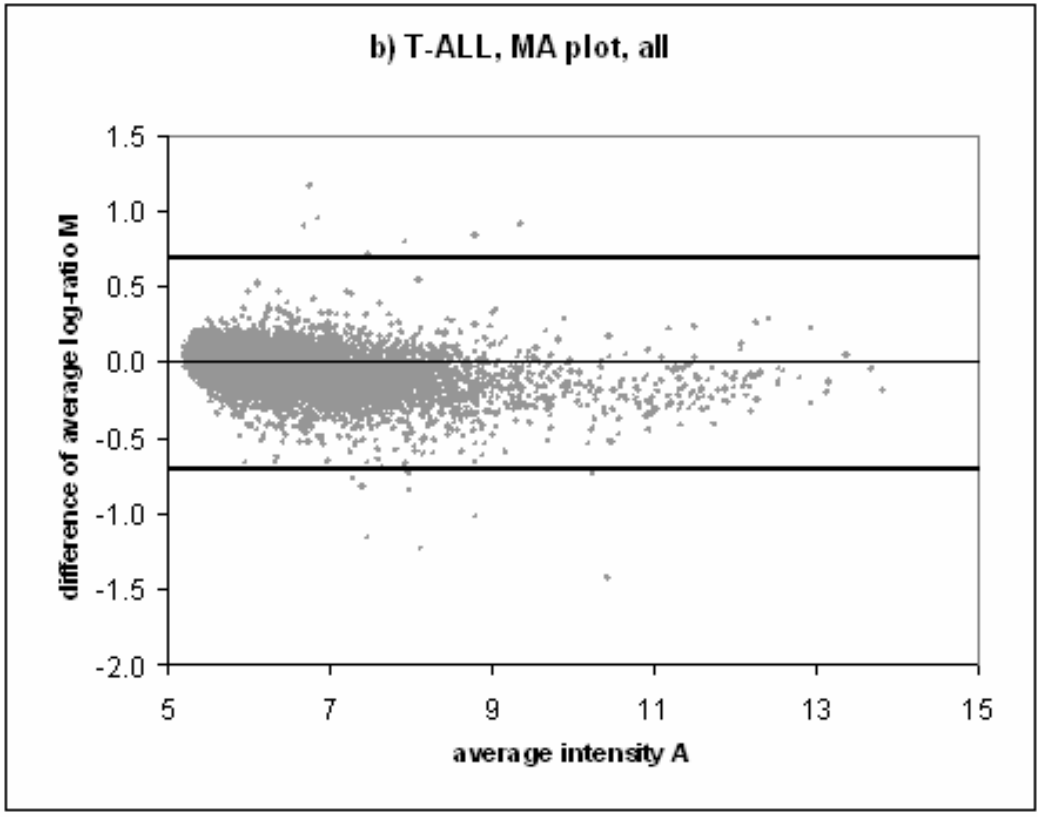
- scatter plot matrix does not distinguish
- density plot suggests poor quality
- supported by percentage of present calls
  - only chip with < 10% present calls

⇨ poor hybridisation, undetected by

- average background, noise level, control spots, etc.
- image inspection (at least not clear)

Two other subgroups showed similar behaviour. Same problem?



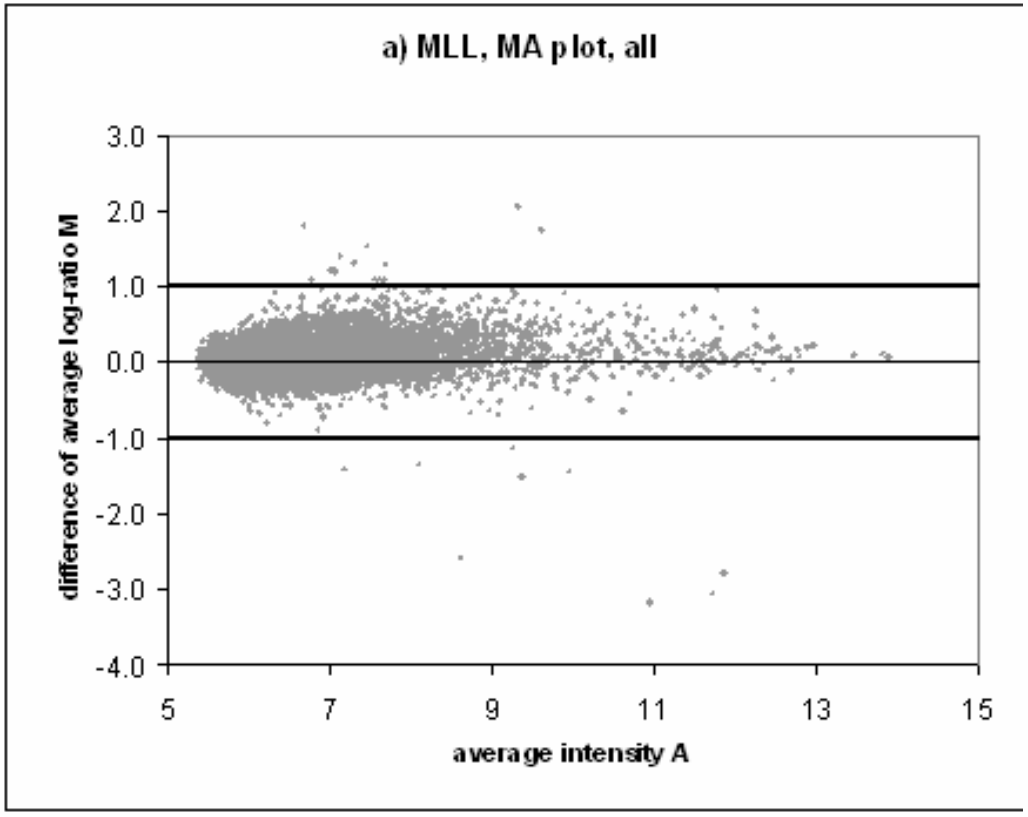


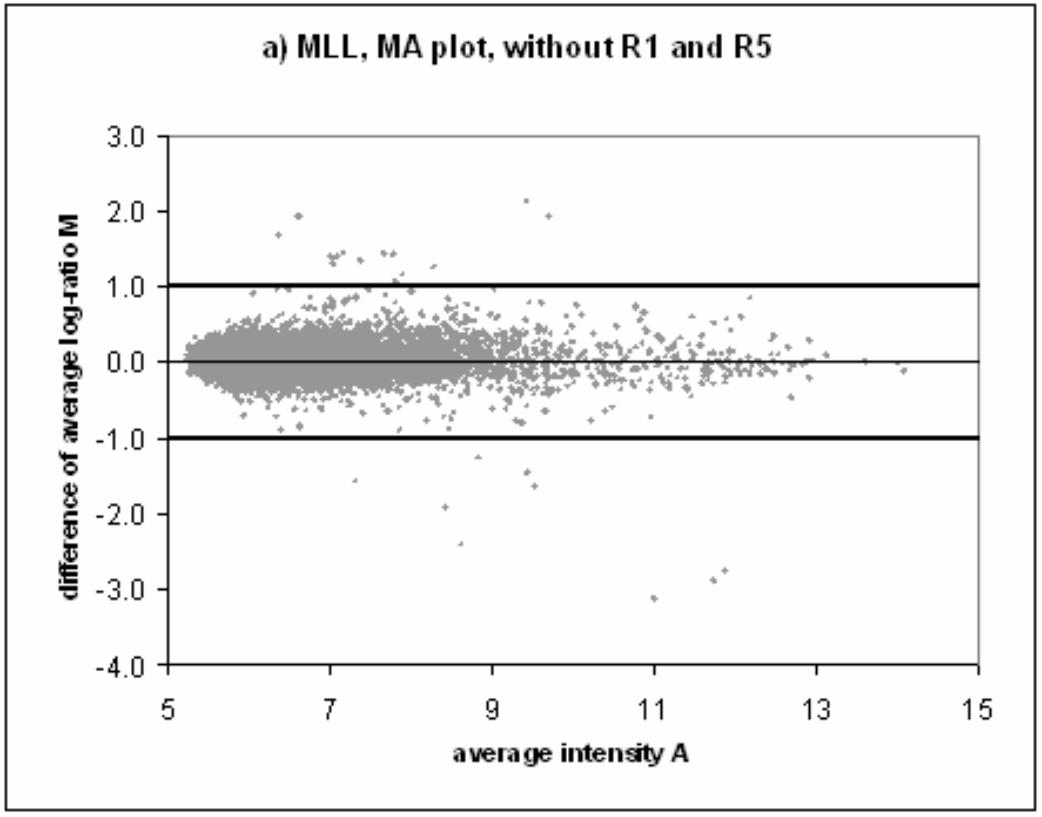
MLL: $n_1=6$, $n_2=5$

T-ALL: $n_1=26$, $n_2=8$

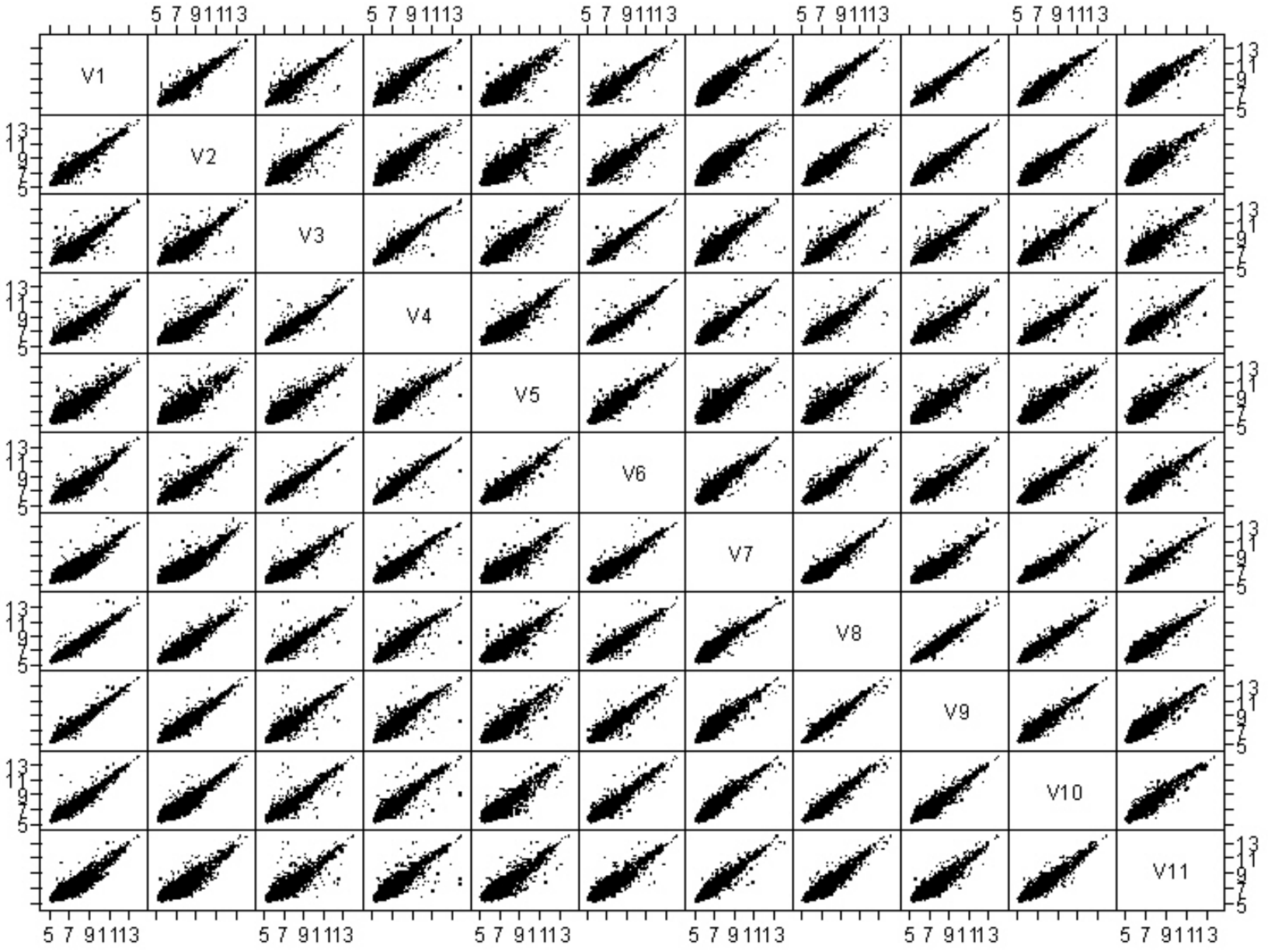
Percentage of present calls shows outliers, no other subset worked

	C1	C2	C3	C4	C5	C6	R1	R2	R3	R4	R5
Noise (RawQ)	13.79	11.65	14.55	14.42	16.70	13.40	25.63	15.34	13.78	20.58	16.48
Background Avg.	429.49	331.63	459.69	419.92	664.96	456.88	645.31	487.35	406.01	526.38	372.53
Background Std.	5.84	10.24	6.09	6.06	25.76	12.58	13.38	5.91	5.23	21.92	5.06
Number Present	25.7%	34.3%	23.4%	15.8%	15.8%	20.7%	7.4%	21.1%	27.2%	16.4%	9.9%
Number Absent	72.3%	63.6%	74.8%	82.2%	83.0%	77.5%	91.2%	76.8%	70.7%	81.8%	88.7%
Number Marginal	2.0%	2.1%	1.8%	2.0%	1.2%	1.8%	1.4%	2.1%	2.1%	1.8%	1.4%





MLL: scatter plot matrix detects no outliers

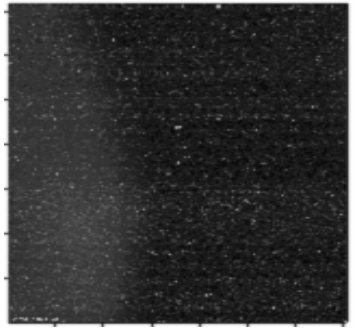


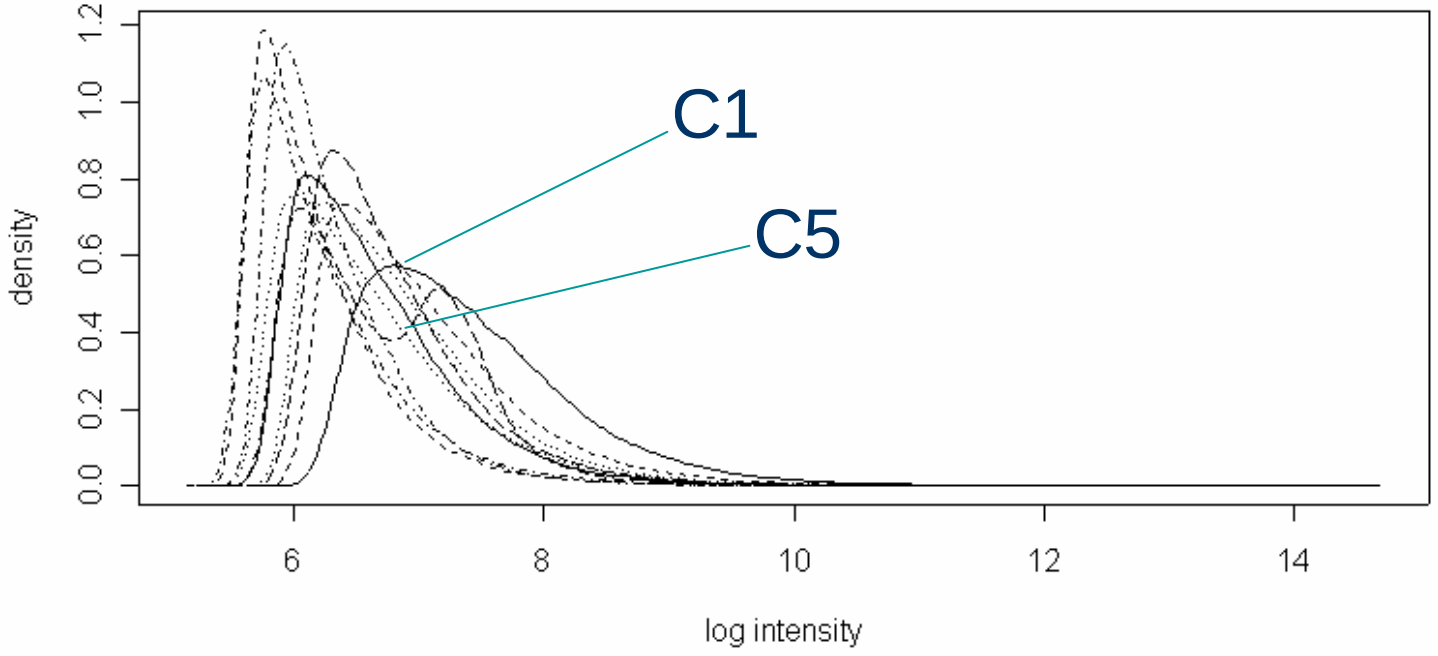
Example 2: present calls and density plots needed

Data: 10/11 samples of MLL subgroup hybridised to HG_U133B

- prior quality filtering
  - percentage of present calls < 10%
  - GAPDH 3'/5' ratio > 3 (housekeeping gene)
- again: MvA plots slightly skewed

C5





Summary: methods for outlier detection are very useful (data not shown)

- scatter plot & correlation matrix
  - good for sample size < 15
  - scatter plots also detect local artefacts
- principal components analysis
  - good for sample size > 15
  - PCA and other methods successfully applied (e.g. Stivers et al, CAMDA 02)
  - unclear if principal components reflect biological or systematic differences

Summary: at least present calls & density plots needed to detect poor quality

- percentage of present calls
  - global test of hybridisation specificity (MMs measure cross-hybridisation)
  - ⇨ MMs do contain useful information
  - from 153 chips 8 with <10% (27 with <15%)
  - threshold: minimal number to obtain good MvA plot (one group had very few replicates)
- density plot of intensities
  - quantile normalisation uses “average” density
  - excluded 2 of 10 (U133, example 2) due to different shape of density estimate

References

Affymetrix: GeneChip® Expression Analysis Technical Manual, Chapter 1 & 2. Updated April 2003. http://www.affymetrix.com/support/technical/manual/expression_manual.affx

Bioconductor Project: http://www.bioconductor.org

Bolstad et al: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, 19 (2003), 185-193.

Ross et al: Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood, (2003), prepublished online May 1.

Yeoh et al: Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell, 1 (2002), 133-143.

Data from Yeoh and Ross: http://www.stjuderesearch.org/data/ALL1 and /ALL3