# Robust Probabilistic Averaging (RPA)

Leo Lahti*
University of Helsinki
leo.lahti@iki.fi

March 15, 2013

# 1 Introduction

*RPA* (Robust Probabilistic Averaging)[1] is a fully scalable preprocessing method for arbitrarily large short oligonucleotide microarray atlases. It outperforms the standard RMA model [1] on moderately sized data sets [2, 3] and unlike the another scalable alternative, fRMA [4], it is applicable to all short oligonucleotide platforms. RPA can be additionally used for probe reliability analysis, and for general summarization of normally distributed multivariate observations that target the same object with varying degrees of noise and bias.

# 2 Usage examples

For usage examples on-line, see https://github.com/antagomir/RPA/wiki

# 3 Methodology and Citations

For details, see [2, 3]. Kindly cite these documents if you find the method useful.

# 4 Details

This document was written using:

```
> sessionInfo()
```

---

*http://www.iki.fi/Leo.Lahti
[1]http://bioconductor.org/packages/release/bioc/html/RPA.html

```
R version 2.15.2 (2012-10-26)
Platform: x86_64-unknown-linux-gnu (64-bit)

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
 [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=C                 LC_NAME=C
 [9] LC_ADDRESS=C               LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] RPA_1.14.0        affydata_1.11.17  affy_1.36.1       Biobase_2.18.0
[5] BiocGenerics_0.4.0

loaded via a namespace (and not attached):
[1] affyio_1.26.0        BiocInstaller_1.8.3  preprocessCore_1.20.0
[4] tools_2.15.2         zlibbioc_1.4.0
```

# References

[1] R. A. Irizarry, B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs, and T. P. Speed. Summaries of Affymetrix GeneChip probe level data. *Nucl. Acids Res.*, 31(4):e15, 2003.

[2] L. Lahti, L. L. Elo, T. Aittokallio, and S. Kaski. Probabilistic analysis of probe reliability in differential gene expression studies with short oligonucleotide arrays. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 8(1):217–225, 2011.

[3] L. Lahti, A. Torrente, L. L. Elo, A. Brazma, J. Rung. A fully scalable online-preprocessing algorithm for short oligonucleotide microarray atlases *Nucl. Acids Res.* 2013. In press. Preprint available at http://arxiv.org/abs/1212.5932

[4] MN McCall and RA Irizarry. Thawing Frozen Robust Multi-array Analysis (fRMA) *BMC Bioinformatics* 12:369, 2011.