DISSERTATIONES
NEUROSCIENTIAE
UNIVERSITATIS
TARTUENSIS
21

STEN ILMJÄRV

Estimating differential expression from multiple indicators

DISSERTATIONES NEUROSCIENTIAE UNIVERSITATIS TARTUENSIS
21

DISSERTATIONES NEUROSCIENTIAE UNIVERSITATIS TARTUENSIS
21

STEN ILMJÄRV

Estimating differential expression from multiple indicators

UNIVERSITY OF TARTU
PRESS
1632

Department of Physiology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia

Dissertation was accepted for the commencement of the degree of Doctor of Philosophy (Neurosciences) on January 22th, 2015 by the Council of the Commencement of Doctoral Degree in Neuroscience

Supervisors:	Hendrik Luuk, PhD, Senior Researcher, Department of Physiology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia Eero Vasar, MD, PhD, Professor, Department of Physiology, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia Jaak Vilo, PhD, Professor, Institute of Computer Science, University of Tartu, Estonia
Reviewers:	Viljar Jaks, MD, PhD, Senior Researcher, Institute of Molecular and Cell Biology, University of Tartu, Estonia Martti Laan, MD, PhD, Senior Researcher, Institute of Biomedicine and Translational Medicine, University of Tartu, Estonia
Opponent:	Djork-Arné Clevert, PhD, Senior Research Fellow, Marie Curie Fellow, Institute of Bioinformatics, Johannes Kepler University of Linz, Austria
Commencement:	April 24, 2015

This research was supported by the European Social Fund's Doctoral Studies Internationalization Program DoRa carried out by Archimedes Foundation, Estonian Science Foundation (grant 9099) and Estonian Research Council (grant PUT120).

ISSN 1736-2792
ISBN 978-9949-32-774-4 (print)
ISBN 978-9949-32-775-1 (pdf)

Tartu Ülikooli Kirjastus
www.tyk.ee

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

1. **Ilmjärv S**, Hundahl CA, Reimets R, Niitsoo M, Kolde R, Vilo J, Vasar E, Luuk H. 2014. Estimating differential expression from multiple indicators. Nucleic Acids Research, 42(8), e72
2. Hundahl CA, Luuk H, **Ilmjärv S**, Falktoft B, Raida Z, Vikesaa J, Friis-Hansen L, Hay-Schmidt A. 2011 Neuroglobin-deficiency exacerbates Hif1A and c-FOS response, but does not affect survival during severe hypoxia *in vivo*. PLos ONE 6(12): e28160
3. **Ilmjärv S**, Reimets R, Hundahl CA, Luuk H. 2014. Effect of light on global gene expression in the Neuroglobin-deficient mouse retina. Biomedical Reports 2(6): 780–786

Contribution of the author:

1. The author co-developed the statistical method behind DEMI, designed and implemented DEMI in R package, performed differential expression analysis with microarray and RNA-seq pipelines used in benchmarking, co-wrote the manuscript.
2. The author performed microarray differential expression analysis.
3. The author performed the microarray differential expression analysis, calculated the enrichment between different datasets and co-wrote the manuscript.

ABBREVIATIONS

A	–	Adenine
AE	–	ArrayExpress
APT	–	Affymetrix Power Tools
AUC	–	Area Under the Curve
C	–	Cytosine
CDF	–	Chip Definition File
cDNA	–	complementary DNA
ChiP-chip	–	Chromatin immunoprecipitation DNA microarray
Cirbp	–	Cold-inducible RNA-binding protein
Cygb	–	Cytoglobin
DEMI	–	Differential Expression from Multiple Indicators
DFW	–	Distribution Free Weighed
DNA	–	Deoxyribonucleic acid
FARMS	–	Factor Analysis for Robust Microarray Summarization
FDR	–	False discovery rate
FNR	–	False negative rate
FPR	–	False positive rate
G	–	Guanine
GEO	–	Gene Expression Omnibus
GO	–	Gene ontology
H3K9ac	–	Histone 3 lysine 9 acetylation
H3K27me3	–	Histone 3 lysine 27 tri-methylation
HBR	–	Human Brain Reference
HIF-1/Hif-1	–	Hypoxia-inducible factor 1
HIF-2/Hif-2	–	Hypoxia-inducible factor 2
Hif1A	–	Hypoxia-inducible factor 1-alpha
Hprt1	–	Hypoxanthine phosphoribosyltransferase 1
ICH	–	Immunohistochemical
IR	–	Immuoreactive
LDTg	–	Laterodorsal tegemental nucleus
LH	–	Lateral hypothalamus
LRES	–	Long-range epigenetic silencing
MAQC	–	MicroArray Quality Control
MCC	–	Matthews correlation coefficient
MEF	–	Mouse embryonic fibroblasts
MIAME	–	Minimum Information About a Microarray Experiment
miRNA	–	microRNA
MM	–	Mismatch probe
mRNA	–	messenger RNA
Ngb	–	Neuroglobin
PLIER	–	Probe Logarithmic Intensity Error
PM	–	Perfect match probe

PPTg	–	Pedunculopontine tegmental nucleus
qPCR	–	quantitative Polymerase Chain Reaction
RMA	–	Robust Multi-Array Average
RNA	–	Ribonucleic acid
RNA-seq	–	RNA-sequencing
ROC	–	Receiver Operating Characteristic
SCN	–	Suprachiasmatic nucleus
T	–	Thymine
TN	–	True negative
TNR	–	True negative rate
TP	–	True positive
TPR	–	True positive rate
UHR	–	Universal Human Reference
wt	–	wild-type

INTRODUCTION

With very few exceptions every cell in our body contains identical copies of the same deoxyribonucleic acid (DNA) molecules. This is in contrast with the vast diversity of cell types that essentially derive from a single genome. To accomplish this variety, distinct nucleotide sequences on the genome called genes are used as templates for ribonucleic acid (RNA) molecule synthesis. This process can vary considerably in its activity from cell type to cell type. By carefully regulating the activity of genes, cells are able to obtain, maintain and change their function and physiology. This property allows cells to adapt to changing environments and to develop specific behavior essential for their function and role in the body in both health and disease. Therefore quantifying gene activity across cell types and diverse conditions is an incredible source of information that can be utilized to understand mechanisms of disease, develop intervention strategies and much more.

The activity of genes is most often characterized by the level of messenger RNA (mRNA) synthesized using the DNA sequence of a gene as the template. By measuring the amount of mRNA product from a single gene, i.e. gene expression level, we are able to determine the transcriptional activity of that gene. There are several technologies that can be used to measure gene expression levels. There are many more analysis methods developed to analyze the data that these technologies produce. Each of them has its peculiarities with unique strengths and weaknesses. Often they approach the data from alternative angles, making divergent assumptions and providing controversial results that can make the methods difficult to compare. Therefore, researchers have the choice on which technology and analysis method would best fit the research question in order to obtain accurate results.

The current thesis addresses these issues and proposes a novel alternative, DEMI (Differential Expression from Multiple Indicators), for analyzing gene expression data produced by high-density microarrays. DEMI takes advantage of the repeated measurements of the target mRNA level by independent probes on the microarrays. Integrating data from multiple probes is somewhat related to the law of large numbers, which states that in sufficiently large samples, estimates about the distribution should converge with the actual value of the parameters in the population. By design, DEMI is versatile and can be used to address various research questions, from a standard comparison of two groups to a more complex time-series analysis, as well as a novel approach to analyze differential expression of genomic regions encompassing neighboring genes. DEMI has been applied to study the effect of Neuroglobin (Ngb) deficiency on the global gene expression in the brain following hypoxia and the effect of light in the mouse retina. We have also used it to study the effects of prolonged hypoxia and hypothermia in cell culture. To demonstrate its accuracy, DEMI was benchmarked in reference to state of the art methods on several datasets with various experimental designs. I will argue that DEMI is easy to use and

simple to understand, provides enough flexibility to accommodate various experimental designs and retains accuracy even with low number of samples and, hence, has a wide application range for both biologists and bioinformaticians alike. Methodological concerns arising during benchmarking as well as future perspectives on extending DEMI to RNA sequencing (RNA-seq) and proteomics are discussed. DEMI software is freely available which has been implemented as an R package.

I hope that, eventually, this work will serve as a good reference for new enthusiasts to dig deeper into the intriguing realm of gene expression data analysis. Hopefully, DEMI will reduce the learning curve and support decision making for individual research objectives. No method should be used as the ultimate source of truth and using several methods together with eyes wide open would maybe best serve the researcher to increase confidence in the results.

REVIEW OF LITERATURE

1. Gene expression measurement technologies

The physiology and function of a living cell is specified by the collaborative activity or inactivity of every single gene in the genome. This activity is carried out by the controlled synthesis of messenger RNA (mRNA) from the template deoxyribonucleic acid (DNA) strand, in a process called transcription (Alberts, 2014).

Gene expression level is the extent of gene's activity that is determined by the amount of mRNA molecules in the cell where higher rates of transcription lead to higher levels of gene expression. Although proteins perform the actual functional tasks, in general, the level of mRNA is considered a reasonably good estimate of the protein level (Schwanhausser et al., 2011). Therefore disturbances altering gene expression can alter cell's physiology and function, making it an important study subject to enhance the understanding of molecular mechanisms at work. For example, by measuring mRNA levels in the cells we can discover properties responsible for the differences between cell types. One can use it, for example, to describe abnormalities in different cancer types (reviewed in van't Veer and Bernards, 2008) or characterize the differentiation stages of engineered neural tissue from reprogrammed somatic cell lines (Wang et al., 2011).

There are several technologies capable of measuring mRNA levels, each with their own strengths and weaknesses as mentioned by Bradford et al., 2010 and by Liu et al., 2011. These technologies could broadly be categorized into single-gene and whole-genome measurements. A single-gene approach, considered as the "gold-standard" by many, is quantitative polymerase chain reaction (qPCR). Currently, the two main technologies to measure gene expression of many genes simultaneously are microarrays (Schena et al., 1995) and sequencing approaches (reviewed in Metzker, 2010) whereas qPCR is often used to confirm these findings (Qin et al., 2006). In the following chapters I will focus on describing the steps involved with differential expression analysis using microarrays.

1.1. Microarrays

1.1.1. Description of microarrays

Double-stranded DNA and RNA molecules exhibit the characteristic pairing of matching nucleotide bases by forming hydrogen bonds. It is known as the complementarity principle and is heavily used by many processes in the cell, e.g. by microRNA (miRNA) to recognize their targets (Brennecke et al., 2005) or to direct meiotic recombination of chromosomal DNA. In double-stranded DNA, adenine (A) is always paired with thymine (T) and cytosine (C) is paired with guanine (G). The complementarity principle is applied in technologies that

utilize hybridization, a process of annealing two complementary sequences together. By measuring the magnitude of hybridization between complementary molecules, researchers can measure the amount of sequence-specific signals, thus making direct observations of gene expression levels. Microarrays were specifically designed for such purpose.

DNA microarrays are small solid surfaces where oligonucleotides are covalently attached and organized into arrays of "spots" (reviewed in Bumgarner, 2013). Each spot contains millions of identical oligonucleotides (termed "probe") representing a small stretch of sequence identical or complementary (depending on the array) to an expressed sequence, the "target". The probes generate a signal when hybridized with fluorescently labeled complementary DNA (cDNA) synthesized from the RNA samples studied. The signal intensity depends on the amount of target RNA in the original samples with higher amounts increasing the chance of hybridization, thus yielding a higher signal from the spot.

Although there are several types of microarrays produced by distinct companies, the current work concentrates on oligonucleotide microarrays produced by the company Affymetrix®, the most popular microarray provider. Compared with other manufactures, Affymetrix® set themselves apart by being the first to introduce microarrays providing repeated measurements of the target with many probes targeting the same mRNA species. For example Human Exon ST 1.0 microarray contains more than 5.5 million probes with length of 25 nucleotides whereas Illumina HumanHT-12 v4 Expression BeadChip contains only 47,231 probes with length of 50 nucleotides (http://www.affymetrix.com/estore/catalog/131452/AFFY/Human+Exon+ST+Array#1_1, last access: 30.01.2015; http://res.illumina.com/documents/products/product_information_sheets/product_info_humanht-12.pdf, last access: 30.01.2015). Although smaller probe size can result in poorer hybridization affinity, the greater number of probes on Affymetrix® arrays allows to interrogate a higher number of targets with more fine-grained resolution. In addition, smaller probe size is less prone to non-specific cross-hybridization as it is easier to design shorter probes measuring unique genomic sequences not shared by homologous genes or gene families (Relogio et al., 2002). High number of probes also allows interrogating one target from several locations. For example, probes can be designed to target individual exons, which can then be used to estimate the use of alternative transcripts (Bemmo et al., 2008; Johnson et al., 2003). Furthermore, one exon can be targeted by several probes (a probe set) increasing the reliability of signal estimates through replicated measurements. In addition, the multiple probes per target design should even the odds by improving specificity and reliability (Miller and Tang, 2009).

Using microarrays for large-scale gene expression studies has been a common practice since the mid 1990s (Schena et al., 1995). Later, Brazma et al. (2001) specified the Minimum Information about Microarray Experiment (MIAME) standard. The goal of the MIAME standard was to ensure that

experiments carried out by the researchers could be independently verified and the results easily interpreted while also promoting the establishment of public repositories and data analysis tools (Brazma et al., 2001). Since the establishment of ArrayExpress (Rustici et al., 2013) it contains over 42,000 microarray experiments in the public archive. Most of these experiments have only been looked at in confinement from a narrow research perspective but this data can also be used for conducting large-scale meta-analysis studies. For example, such approaches have been used to study individual cancers from several resources as well as across different cancer types, which decreases the limitations from biological and technical biases and small sample size (reviewed in Chen et al., 2014). Upon that bioinformaticians have built added value databases (Rung and Brazma, 2013) that, for example, can investigate the coexpression of genes over a large number of datasets (Adler et al., 2009) or allow researchers to query conditions where specific genes are differentially expressed (Chen et al., 2008).

1.1.2. Microarray weaknesses and strengths

A notable limitation of microarrays is its fixed design (Liu et al., 2011; Zhao et al., 2014). The probe sequences are selected by the manufacturer based on knowledge of gene annotations at the time of design. With reference genomes being consistently updated and genetic information improved, the arrays become outdated as new coding regions are discovered and the annotation of known genes is refined. However, by updating array designs according to improved probe annotations one can discard probes with poor hybridization properties and update probe sets to target novel genes.

Another drawback of microarrays is the fixed number of oligonucleotide copies of each probe on the array. It can lead to under-estimation of the true mRNA abundance if there are more labeled target molecules than complementary oligonucleotides to bind them to the array, which can yield in a saturated signal (Wang et al., 2009). Therefore, the fluorescence signal can reach a plateau when excess mRNA molecules have no space for hybridization ("ceiling effect"). This presents a limit for the dynamic range of expression signals on the microarrays. In addition, the detection of low abundance mRNAs is complicated by the background noise and non-specific hybridization, which make it difficult to distinguish between "no" and "low" expressed genes (Zhao et al., 2014).

Oligonucleotides varying in their sequence have alternate hybridization properties (Zhao et al., 2014). Some have stronger affinity towards the target mRNA than others. This makes it difficult to compare different probes in a probe set directly because signal intensity does not always correspond to the actual amount of mRNA (Zhao et al., 2014). This can present a problem during microarray analysis where summarization procedure, discussed later, combines the signal from relevant probes into a probe set value. In addition, several probe

sets can correspond to the same gene and it is not uncommon that they do not yield a consensus (Zhao et al., 2014).

Despite well-known issues with microarrays, the benefits of using microarrays include relatively low cost, high throughput and well-established analysis workflows. Therefore microarrays are still the tool of choice in experiments where many samples are available (Rung and Brazma, 2013).

2. Differential expression analysis with microarrays

Quantifying the expression of genes across cell types and diverse conditions is an incredible source of information that can be utilized to understand mechanisms of disease, develop intervention strategies and much more. It is therefore important to design the experiment properly for it is impractical to study cancer cells without studying healthy cells when we aim to capture the difference. In most cases current gene expression measurement technologies and downstream analysis methods don't perform direct quantification of gene expression levels. Instead, they provide relative estimates and compare them between experimental treatments. Differential expression is typically defined as a statistically significant difference in the abundance of an mRNA species between two treatments and it can be used to identify genes with altered expression levels (Slonim, 2002). Differential expression analysis is the most common application of microarray data (Shedden et al., 2005).

In the simplest cases researchers perform differential expression analysis between two groups, typically labeled as "case" and "control". However, differential expression analysis can also be applied to other experimental designs, for example time-series, concentration series or factorial designs. Experimental designs, in principle, are limited only by the researcher's imagination and the existence of applicable statistical tools.

2.1. Preprocessing of Affymetrix® microarrays

There are several sources that contribute to signal variation between microarrays. In addition to the *interesting variation*, that depends on the genetic and environmental conditions (Hartemink et al., 2001), there is also *obscuring variation*. Sources for obscuring variation derive from array manufacturing, processing (hybridization, washing and scanning), sample preparation (RNA extraction and labeling) and sloppy experimental design (Irizarry et al., 2003b). Therefore, in order to make signal intensities on different arrays comparable, we need to preprocess the data, which essentially projects the intensities from different arrays onto a common scale.

Preprocessing of Affymetrix® microarrays consists of several independent steps. Often, the first step is to correct the raw signal intensities for the background noise, which on the microarrays is a signal caused by optical noise

during the scanning process and non-specific binding (Irizarry et al., 2003b). Secondly, the background-corrected signal intensities have to be normalized so that they would have the same signal distribution across all arrays in the study. This serves to reduce the problem of obscuring variation (Bolstad et al., 2003). Finally, as the probe level data by itself is uninformative it needs to be aggregated into a single expression value ultimately representing a gene or transcript, which can then be used to evaluate differential expression.

In the following chapters I will describe four different preprocessing methods and the steps they involve. These methods are Robust Multi-array Average, RMA (Irizarry et al., 2003b), Factor Analysis for Robust Microarray Summarization, FARMS (Hochreiter et al., 2006), Distribution Free Weighted, DFW (Chen et al., 2007) and Probe Logarithmic Intensity Error, PLIER (Affymetrix, 2005). In addition, I will give an overview of two approaches to estimate differential expression. The first one uses a mixture of linear modeling and empirical Bayes and is implemented in the Limma package (Smyth, 2004) and the second one, RankProd (Hong et al., 2006), uses a non-parametric rank product approach. In the later chapters I will introduce a novel differential expression analysis method called DEMI (Differential Expression from Multiple Indicators), which will be compared against different combinations of previously mentioned methods.

Although some methods like DFW or FARMS were specifically developed to tackle the summarization problem, these packages also incorporate preceding preprocessing steps – e.g. quantile normalization developed for RMA is also used in FARMS.

2.1.1. Background correction

Signals on individual arrays are always affected by variations not related to the experimental factors. As mentioned before, this variation is called obscuring variation that when ignored can lead to inaccurate results (Irizarry et al., 2003b). At high target abundance, the contribution of background noise to signal is proportionally smaller since most of the spot intensity is target-derived whereas for targets with low observed intensities the background signal is proportionally larger (Affymetrix, 2005). In other words, background noise has a substantial effect on the intensity values, especially when the intensities are low (Chen et al., 2007). The presence of background noise is obvious since the minimum probe intensity is never 0 (Zhijin et al., 2004) as you would expect in an actual biological situation where a gene is expressed below detection level. Therefore it is often recommended to adjust the data for this variation using background correction.

From the previously mentioned methods RMA is probably the most widely used open-source analysis method for Affymetrix® microarrays with almost 6000 citations as of July 2014. It was developed at a time when Affymetrix® microarrays were designed to take advantage of mismatch (MM) probes to

estimate non-specific hybridization. Unlike perfect match (PM) probes, which are perfectly complementary to the target, MMs have a mismatching nucleotide in the middle (at the 13th position) of the 25-mer sequence. MMs were incorporated on arrays to enable the subtraction of background noise by subtracting MM signal from the paired PM signal, which initially was the default approach for background noise adjustment (Zhijin et al., 2004). Different to previous understanding, RMA developers ignored MMs because they claimed that although non-specific binding is contained in the MM values, empirical results demonstrate that mathematical subtraction does not translate into biological subtraction (Irizarry et al., 2003a) and hence only PM values should be used. Similar to RMA, both the developers of FARMS and DFW have chosen to ignore MM values in their software packages and use only PM values to calculate gene expression estimates. Most likely as a response to this trend, the newer generation Affymetrix® microarrays do not contain mismatch probes.

Not all methods utilize background correction. From the discussed methods only RMA and PLIER perform background correction whereas FARMS, DFW and DEMI start by normalizing the data. In the case of RMA, background-corrected PM values are calculated for every microarray separately. Authors of RMA assume that the signal observed is a convolution of exponentially distributed true signal and normally distributed background noise (Irizarry et al., 2003b). RMA incorporates kernel density estimation for adjusting PM probe intensities. As an alternative to RMA, Affymetrix® developed PLIER. It can be found in the Affymetrix Power Tools (APT) Software Package (http://www.affymetrix.com/estore/partners_programs/programs/developer/tools/powertools.affx, last access: 30.01.2015) as a command line program called *apt-probeset-summarize*. This function allows the researcher to select the appropriate background to be used for background correction. The choices include MMs as background or global uniform background such as genomic or anti-genomic probes that are designed to exclude sequences from any known transcript. In the cases where the researcher assumes the background noise to be minimal only PM values can be utilized or when background is irrelevant to the target response MM values can be added to the PM values (Affymetrix, 2005). However not all arrays have MM probes or dedicated global background probes, so the choices are often dictated by the array design.

After background correction the raw intensity values are replaced with background-corrected signal intensities.

2.1.2. Normalization

The hybridization process is never 100% efficient. Even if the same sample is used on several arrays slight variations in signal intensity distributions can be observed between the arrays (Gautier et al., 2004). To eliminate these inter-array differences it is required to normalize the signal intensities making individual microarrays comparable to each other.

Several normalization procedures have been proposed by different authors of which the most popular is *quantile normalization*, which was first used by the authors of RMA and described in Bolstad et al. (2003). In their description of quantile normalization, signals are first sorted according to the intensity. The individual probe intensity values at every quantile are then replaced with the average probe intensity values of that quantile after which the signals are rearranged back to the initial ordering. Although the arrays are now comparable to each other quantile normalization forces the values of quantiles across all arrays to be equal and therefore risks removing some of the signal in the tails (Bolstad et al., 2003). However this problem should be diminished through the use of multiple probes in a probe-set during summarization process (Bolstad et al., 2003).

FARMS and DFW, which were developed as alternative summarization methods also utilize quantile normalization, as does PLIER. FARMS users have the additional option of using cyclic loess normalization. The latter implements the M versus A plot that was initially developed for two-color microarrays by Dudoit et al. (2002). Bolstad et al. (2003) modified it for single-channel micro-arrays where instead of two colors they used two separate arrays to calculate logarithmic difference of expression and logarithmic expression mean for M and A values respectively. They then used these values to fit a normalization curve with a local regression method called loess for each pairwise combination of the arrays.

There are many other normalization methods, not discussed in this work, available for the researchers. For example, solely the *affy* package (Gautier et al., 2004) available in BioConductor (Gentleman et al., 2004) includes *quantiles*/*quantiles.robust*, *loess*, *contrasts*, *constant*, *invariantset* and *qspline*. Comparing cyclic loess and quantile normalization Bolstad et al. (2003) argues that the practical choice to be used in normalization step would be quantile normalization, especially in terms of speed. Quantile normalization is also the default approach for all the methods discussed in this work apart from DEMI.

2.1.3. Summarization

High-density microarrays are designed to interrogate one target sequence by several probes. Even after background correction and normalization have been performed it remains challenging to make conclusive inference about the expression of target sequence since different probes measuring a single target can have different hybridization kinetics and can produce different expression profiles (Zhao et al., 2014). Therefore, it is necessary to summarize the probe level intensity values of the specified target sequence into a single value represented by a probe set. Considering the methods discussed in this work, summarization is also the primary step of preprocessing that sets them apart. As such, FARMS and DFW were specifically developed as alternative summarization methods.

Summarization is most often performed after the raw signal intensities have been background corrected and normalized. However, in the case of PLIER, normalization can also be applied after summarization (Affymetrix, 2005). In addition, PLIER performs summarization without log-transformation whereas other methods use log 2 transformed signal intensities (probably due to the fact that the distribution of log-transformed data is approximately normal (Sartor et al., 2003). PLIER models the measured signal intensity as the product of the amount of mRNA (target response), common across intensities in a given probe set, and probe's affinity (feature response), common across all arrays in the experiment (Affymetrix, 2005). Initial estimates of the probe set values are calculated using median polish, initially used by the authors of RMA (Affymetrix, 2005; Irizarry et al., 2003b). Median polish is an iterative method operating on a matrix by alternatively extracting row and column medians. Iterations continues until convergence or a limit number of iterations has been reached (Holder et al., 2001). Upon receiving the first estimations PLIER then attempts to find estimates that more precisely fit the data by calculating approximate likelihood minimizing the PLIER function (Affymetrix, 2005).

The authors of FARMS (Hochreiter et al., 2006) proposed a Bayesian factor analysis technique for probe set summarization where they model initial mRNA concentrations from log-transformed PM values. FARMS' main idea is to decompose the covariance structure data of multiple PMs, that measure the same transcript, into a Gaussian distributed latent factor and additive independent noise via the latent variable posterior distribution providing a rather unique ability to assess whether a probe set is informative or not. FARMS initializes by normalizing perfect match observations to a zero mean after which expectation-maximization algorithm, modified to maximize the Bayesian posterior, estimates the model parameters given the data. The last step recovers the true signal from the estimated log-transformed RNA concentrations in the hybridization mix (Hochreiter et al., 2006).

The last method discussed is DFW by Chen at al. (2007). The authors of DFW claim that the model-based methods, such as RMA and FARMS, are heavily dependent on the assumptions made by the model, and require estimation of model parameters in order to perform effectively (Chen et al., 2007). With such reasoning they developed a non-parametric summarization technique where they do not make any distributional assumptions of the probe-level intensities (Chen et al., 2007). Similar to RMA and FARMS, DFW uses log 2 transformed PM intensities. DFW estimates the quality of each PM probe in the probe set by assigning a weight, using Tukey weight function, to each probe. The weight is calculated based on the range of probe signal intensities across all arrays in the analysis. Individual probes in a probe set are then combined by taking account the weighed probe intensities together with range and standard deviation that is calculated on the weighted intensities.

2.2. Differential expression statistics

Several statistical methods have been developed to estimate the statistical significance of differentially expressed genes. As such, Limma (Smyth, 2005) is an extensively used R package that contains methods designed to estimate differential expression by taking preprocessed expression data as input. Limma fits a linear model for every probe set in the expression matrix where it estimates the least square error by calculating coefficients for every group in the experimental design. Using a *contrast matrix* the user can specify which comparisons to extract from the fit. Based on the contrast matrix, standard errors and linear coefficients will be calculated relevant to the comparisons of biological interest for the researcher. This allows the user to conduct any comparison or combination of comparisons. Limma estimates differential expression with a moderated t-test where the standard errors of coefficients have been moderated across all probe sets using an empirical Bayes approach (Smyth, 2005). This can be thought of as condensing the estimated sample variances towards a pooled estimate, which allows to make reliable inferences even when the sample size is small (Smyth, 2004).

RankProd (Hong et al., 2006) is another method used to determine the statistical significance of differential expression on preprocessed microarray data. RankProd provides a non-parametric method that detects genes that are consistently ranked high in a number of gene lists, namely genes that are often found to be strongly up- or downregulated in replicate experiments (Hong et al., 2006). In the first step RankProd computes pair-wise ratios of fold change for all samples between treatment versus control. Secondly, the ratios are ranked within each comparison and the rank product for each gene is calculated. After data permutation the false discovery rate (FDR) associated with each gene can be determined that indicates the significance of differential expression for up- and downregulated genes separately (Hong et al., 2006).

It is in the nature of microarray data that a large number of genes are tested simultaneously resulting in a multiple testing problem. RankProd approaches that problem by estimating the proportion of false positives by permuting gene expression values within each single array. In the case of Limma, the researcher can manually determine which multiple correction approach to use. The default p-value adjustment procedure is the false discovery rate (FDR) by Benjamini and Hochberg (1995) to control for the expected proportion of falsely rejected null hypotheses.

3. The function of Neuroglobin in the brain

The gene *NGB* encodes a protein called Neuroglobin (Ngb), which is a neuron-specific globin that is predominantly expressed in the brain (Burmester et al., 2000). More specifically, it has been demonstrated that the expression of Ngb occurs mainly in the brain regions that are related to functions such as sleep-

wake activities, circadian rythm and regulation of energy homeostasis (Hundahl et al., 2008a; Hundahl et al., 2008b).

The function of Ngb in the brain is subject to controversy. Initially, the function of Ngb was associated with oxygen storage and intracellular transport (reviewed in Brunori and Vallone, 2007) and, in addition, it was shown to possess a neuroprotective role under hypoxic conditions where there is a shortage of oxygen supply to the brain (Fordel et al., 2004; Sun et al., 2001). Contradictorily, however, evidence has been provided that shows a low oxygen affinity of Ngb (Fago et al., 2004). In addition, the low concentration levels of Ngb in neurons question its role as oxygen reservoir and in oxygen transport (reviewed in Brittain et al., 2010).

Despite the obvious disagreements in the function of Ngb, Schmidt et al. (2003) demonstrated that the amount of Ngb protein present in the murine retina is 50 to 100 fold higher in comparison to the total brain extracts making up 4 percent of total retinal protein. The retina is like the brain, one of the most metabolically active tissues (Ames, 1992; Anderson and Saltzman, 1964). To be more specific, oxygen consumption rate in retina is dependent on the exposure to light where dark-adapted retina requires more oxygen compared to light-adapted retina (Ahmed et al., 1993). Therefore, it is reasonable to expect that if Ngb is integral for correct retinal functioning, Ngb-deficiency would have detectable effects on light-induced gene expression if it were important for oxygen metabolism or its delivery.

AIMS OF STUDY

1. Describe and benchmark a novel probe-level differential gene expression analysis workflow (DEMI). Demonstrate the applicability of DEMI to non-conventional experimental designs and different target categories (e.g. gene, transcript, genomic region).
2. Study the effect of acute and prolonged hypoxia on neuronal viability and gene expression on Neuroglobin deficient mice with DEMI.
3. Study the effect of Neuroglobin deficiency on light-induced gene expression response in the mouse retina with DEMI.

MATERIALS AND METHODS

1. DEMI algorithm (Paper I)

DEMI is a probe-level microarray analysis workflow that consists of three steps: I) normalization of the raw signal intensities, II) evaluation of probe-level signal dynamics, III) estimation of differential expression of the target in question. CEL files are read as input and an expression matrix containing raw signal intensities is created where rows represent independent measurements (equivalent to probes) and columns represent individual samples of the experiment.

In contrast to the popular preprocessing methods that use quantile normalization, DEMI utilizes relative rank normalization instead. Relative rank provides an intuitive measure of gene expression by relating the intensity of a probe signal to all other probes on the array. The appointed values will be on the scale of 0 to 100 where a relative rank of 75 corresponds to a probe with signal intensity at 75th percentile (the third quartile) in the probe signal distribution. For clarification, relative rank 0 and relative rank 100 correspond to the weakest and strongest signal respectively. Ranking is applied to each column in the expression matrix separately and is therefore independent of the other samples in the experiment.

Let us consider a simple experiment where we have samples corresponding to two groups or conditions, a TEST and a REFERENCE group. Given a dataset of m TEST and n REFERENCE samples, each represented by q independent measurements on every array, we can construct an expression matrix $\boldsymbol{X} = (x_{ij})$ where $i = 1 \cdots q$ and $j = 1 \cdots m + n$. Relative rank normalization is applied to each column vector R_{*i} yielding a normalized expression matrix $\boldsymbol{X^{norm}}$ with concatenation operation $\|$ of all column vectors where the order of rows and columns remains identical to the original expression matrix.

$$\boldsymbol{X^{norm}} = \frac{1}{q} \times (R_{*1} \parallel R_{*2} \parallel \cdots \parallel R_{*m+n})$$

In the second step DEMI evaluates whether the experimental treatment has a statistically significant effect on the normalized signal intensities. This process is applied to each probe separately in the normalized expression matrix $\boldsymbol{X^{norm}}$. The choice of test method depends on the research question. Among the choices are, for example, Wilcoxon-Mann-Whitney rank sum test to compare two groups of samples and Kendall's tau statistic to evaluate departure from monotonicity. The Wilcoxon-Mann-Whitney rank sum test is the default choice in DEMI when estimating differential expression between TEST and REFERENCE samples.

Given a set of probes $\boldsymbol{Q} = \{q_1, \cdots, q_q\}$, a normalized expression matrix $\boldsymbol{X_i^{norm}}$ where i specifies the current row, and sets of $\boldsymbol{M} = \{m_1, \cdots, m_m\}$ and $\boldsymbol{O} = \{o_1, \cdots, o_n\}$ that correspond to the column indices indicating TEST and REFERENCE samples, respectively, we will define a set of upregulated probes $\boldsymbol{H}$ and a set of downregulated probes $\boldsymbol{L}$. Depending on the sample size these sets $\boldsymbol{H}$ and $\boldsymbol{L}$ are constructed differently. For all tests where both groups have more than 3 samples ($m, n > 3$) a probability of null hypothesis (H0) being true, i.e. there is no difference between the TEST and REFERENCE group for probe q_i, is calculated by obtaining a sum of ranks that is higher or lower than the observed rank sum of the TEST samples. Upregulated and downregulated sets of probes can be formulated as

$$\boldsymbol{H} = \{q_i : if\ P_{higher}(r_i) < 0.05\}$$

$$\boldsymbol{L} = \{q_i : if\ P_{lower}(r_i) < 0.05\}$$

where r_i is a vector of ranks corresponding to probe i and P_{higher} and P_{lower} corresponds to the H0 probability of obtaining a sum of ranks higher or lower, respectively, than the observed rank sum of TEST. In the cases where either the TEST or REFERENCE group contains less than 4 samples ($m \leq 3$ or $n \leq 3$) we use a heuristic where a probe is labeled differentially expressed only if all TEST ranks are either higher or lower than corresponding REFERENCE ranks.

Similar to the summarization problem of other methods, probe-level differential expression information by itself is uninformative and needs to be combined into a target level differential expression estimate. Following our example we want to test whether the target (e.g. gene) is differentially expressed between TEST and REFERENCE samples. Given a set of targets $\boldsymbol{T} = \{t_1, \cdots, t_t\}$ where each target t_i is related to a predetermined set of probes $\boldsymbol{Q_{t_i}} \in \boldsymbol{Q}$ and given sets of distinct probe expression profiles $\boldsymbol{H}, \boldsymbol{L} \in \boldsymbol{Q}$ that were calculated in the second step, two complementary differential expression estimates will be obtained (one for upregulation and the other for downregulation) using hypergeometric probability distribution. Hypergeometric probability function (P_{hg}) estimates for each t_i the H0 probability of observing $\left|\boldsymbol{Q_{t_i}} \cap \boldsymbol{H}\right|$ or $\left|\boldsymbol{Q_{t_i}} \cap \boldsymbol{L}\right|$ target-specific differentially expressed probes. The H0 probability of target t_i being upregulated ($P_{higher}(t_i)$) and downregulated ($P_{lower}(t_i)$) is formulated as:

$$P_{higher}(t_i) = \sum_{k=|\boldsymbol{Q}_{t_i} \cap \boldsymbol{H}|}^{|\boldsymbol{Q}_{t_i}|} P_{hg}\left(k, |\boldsymbol{H}|, |\boldsymbol{Q} \setminus \boldsymbol{H}|, |\boldsymbol{Q}_{t_i}|\right)$$

$$P_{lower}(t_i) = \sum_{k=\left|Q_{t_i} \cap L\right|}^{\left|Q_{t_i}\right|} P_{hg}\left(k, |L|, |Q \setminus L|, \left|Q_{t_i}\right|\right)$$

When exon is the target, off-target probes are defined as the difference between the set of probes specific to the gene containing exon t_i and probes targeting exon t_i. A potential caveat of using the hypergeometric test to estimate differential expression is assuming an H0-distribution where probe signals are uncorrelated between unrelated targets. We will argue, however, that substantial deviations from this assumption would manifest in the ratio of differentially expressed off-target probes and, given a reasonable number of target-specific probes, it will not result in increased false-positive rate.

As differential expression is evaluated for thousands of targets the estimates must be corrected for multiple testing. For target categories, where targets are expected to share many probes (alternative transcripts, overlapping genomic regions) FDR under dependency is used (Benjamini and Yekutieli, 2001). For gene-specific estimates, the original formulation of the false discovery rate procedure is employed (Benjamini and Hochberg, 1995).

2. DEMI R package (Papers I–III)

In this chapter I will discuss the implementation of DEMI algorithm in R and a few design concepts of the software. The software itself is available as an R package called *demi* which can be installed through CRAN or downloaded directly from DEMI website *http://biit.cs.ut.ee/demi* (last access: 30.01.2015).

The implementation of DEMI in R consists of several packages with *demi* package containing the algorithm. In addition, there are data packages that contain probe and target information for different microarrays. These data packages are required to run the analysis since *demi* does not use traditional CDF (chip definition file) to relate probes to targets as most open-source microarray analysis software solutions do. Keeping the algorithm's implementation and data packages separate from each other reduces the need to update them concurrently. The data packages are downloaded and installed automatically when DEMI is run for the first time.

When running the analysis, the first step is performed by the *DEMIExperiment* function, which loads signal intensities from CEL files and gene annotation informations into a *DEMIExperiment* object followed by relative rank normalization. At this point the experimental design is not important since all the microarrays will be normalized separately and one can have as many samples from as many groups as long as all the arrays represent the same platform. In addition, *DEMIExperiment* function allows the researcher to specify a custom function for the normalization of the input data. When specifying parameters for *DEMIExperiment,* the researcher has to select the target category, from a selection of gene, transcript, exon or a genomic region.

In the second step, by default the differential expression on the probe level is estimated with Wilcoxon-Mann-Whitney rank sum test. Using the function *DEMIClust* the user can specify the names of the two groups that are tested for differential expression. It is important to note that these names have to occur in the names of the CEL files or the system will be unable to identify which columns of the expression matrix correspond to which experimental groups. It is always a good practice for a data file to contain as much information on the metadata in its name as possible. Similar to the function *DEMIExperiment*, *DEMIClust* also allows the user to use a custom method to estimate differential expression on probe level. *DEMIClust* object will hold information on probes that were differentially expressed in the specified experimental design.

DEMIDiff function estimates differential expression on the target level by utilizing hypergeometric distribution function as described above. The returned *DEMIDiff* object contains the results of DEMI analysis in a table where targets are sorted according to increasing FDR value. An example of an output table can be seen in Appendix 1.

A more detailed manual that contains description on different functions and its parameters can be found at http://biit.cs.ut.ee/demi/doc/demi.pdf (last access: 30.01.2015).

3. Animals (Papers II, III)

3.1. Neuroglobin knock-out mice (Papers II, III)

For Papers II and III we used Neuroglobin (*Ngb*) knock-out mice (Ngb-null) that were created by genOway (Lyon, France) under the project no. genOway/SST/HSA1-Ngb/260307. In Ngb-null mice, the exons 2 and 3 were floxed and ultimately deleted from the genome with Cre recombinase during the early stages of embryonic development. Specific details on the creation of Ngb-null mice can be found in Paper II.

3.2. Hypoxia model (Paper II)

In Paper II a total of 46 female and 31 male mice, 16–18 weeks old, were exposed to hypoxic conditions where the oxygen level was lowered to 7% in rigid plastic cages with a glass cover (hypoxia chambers). The exposure to hypoxia was limited either to 90 minutes, 2 hours, 24 hours, 48 hours, 72 or 96 hours. However for microarray and qPCR analysis only female mice were used to study gene expression differences between normoxic mice and mice exposed to hypoxia for 90 minutes and 24 hours. At naive conditions 3 normoxic wild-type (wt) and 5 normoxic Ngb-null, after 90 minutes of hypoxia 5 mice from both genotypes and at 24 hours 5 wt hypoxic mice and 4 Ngb-null hypoxic mice were euthanized.

All animal experiments were performed in accordance to the rules of Dyreforsoegstilsynet, Ministry of Justice, Denmark, who approved the studies by issuing the license number 2010/561-1834.

3.2.1. Neuronal survival under hypoxia

Neuronal survival in Ngb-null and wt mice was estimated by immunohistochemical (ICH) staining with Ngb co-localization markers. In lateral hypothalamus (LH) Orexin-A was used and in two hindbrain nuclei, the pendunculopontine tegmental nucleus (PPTg) and the laterodorsal tegmental nucleus (LDTg) Cytoglobin (Cygb) was used. The number of Orexin-A-immunoreactive (IR) and Cygb-IR neurons and IR double-stained for Ngb was counted in wt (n = 3–6) and Ngb-null mice (n = 4–6) during normoxia and after 24 or 48 hours of hypoxia. The cells were counted in three levels of LH and two levels of PPTg and LDTg, 160 μm apart, from digital photomicrographs using ImageJ (http://rsbweb.nih.gov/ij/, last access: 30.01.2015) cell counter plugin as described in (Hundahl et al., 2010). Statistical analysis was performed with Mann Whitney test.

3.3. Retina model (Paper III)

15 wt C57Bl and 15 Ngb-null mice, 12–16 weeks old, were used in the experiment to study the effect of Ngb deficiency in mouse retina upon exposure to light. To that end, the animals were kept in a 12/12 hour light/dark cycle with *ad libitum* access to food and water. Five mice of both genotypes were euthanized by decapitation either in darkness (t0) or after exposure to light for 1.5 or 5 hours.

Animal care and all experimental procedures were conducted in accordance to the principles of Laboratory Animal Care (Law of Animal Experiments in Denmark, publication 1306, November 23, 2007) and approved by the Faculty of Health, University of Copenhagen (Copenhagen, Denmark).

4. Cell culture (Paper I)

4.1. Primary cell culture

Around 1 million primary mouse embryonic fibroblasts (MEF) (Millipore) were seeded onto 100-mm culture dishes and were grown in DMEM (high glucose 4.5g/l with additional 10% FBS and L-glutamine, PAA) in normal conditions (atmospheric oxygen, 5% CO_2 at 37°C) until 60–70% confluent.

4.2. Modeling hypoxia

Hypoxia was initiated on primary cell culture by lowering the oxygen concentration to 1% in multi-gas incubator (Sanyo). Five replicate samples were used at each experimental condition. RNA was extracted after 24 hours from hypoxic and normoxic cells with Trizol® (Life) after which RNA was used for large scale gene expression profiling performed with Mouse Exon 1.0 ST microarrays (Affymetrix®) as instructed by the manufacturer.

4.3. Modeling hypothermia

Hypothermia was initiated by lowering the temperature of the cell culture incubator to 32°C whereas the control batches were incubated at 37°C. Three dishes per group were incubated for various time periods (0h, 0.5h, 1h, 2h, 4h, 8h and 18h). RNA was extract with Trizol® (Life) and the three replicates at each experimental condition were pooled together. Global gene-expression profiling was performed using Mouse Gene 1.0 ST arrays (Affymetrix®) according to the manufacturer's instructions.

5. Microarray data (Papers I–III)

5.1. Microarray processing (Papers I–III)

In all the papers, microarray processing was done by Rigshospitalet Microarray Center, Copenhagen, Denmark (www.rhmicroarray.com, last access: 30.01.2015) with slight variations, mainly differing by the array platforms used in the analysis. In Paper I, we used Mouse Gene 1.0 ST arrays for the study of hypothermia in primary cell-culture model, and Mouse Exon 1.0 ST arrays to describe changes in gene expression induced by hypoxia. In Paper II, described below, we utilized Mouse Gene 1.0 ST arrays but in Paper III Mouse Exon 1.0 ST arrays were used to study the effect of light on Ngb-null mice. The raw data stored in CEL files can be accessed from public repositories such as Gene Expression Omnibus (GEO) or ArrayExpress (AE). The accession numbers are presented in Table 1 in addition to references to the original publications where further details about each microarray processing procedure can be found.

As an example, in Paper II we studied the effect of hypoxia on wt and Ngb-null mice. For that we used 50ng of total RNA from each sample that was amplified using the WT-Ovation Pico RNA Amplification System (Nugen®, San Carlos, CA, USA) according to manufacturer's instructions. Double-stranded cDNA was generated using the WT-Ovation Exon module followed by biotin labeling with the FL-Ovation cDNA Biotin Module V2. The labeled samples were hybridized to the Mouse Gene 1.0 ST GeneChip® array (Affymetrix®, Santa Clara, CA, USA) after which they were washed and stained with phycoerytrin conjugated streptavidin (SAPE) using Affymetrix Fluidics

Station® 450. Fluorescent images were created with Affymetrix GeneArray® by scanning the arrays as described in the Affymetrix GeneChip® instructions. In all the papers, CEL intensity files were generated in the GeneChip® Command Console® Software (AGCC) (Affymetrix®, Santa Clara, CA, USA).

Table 1. Accession numbers and references for datasets generated in Papers I, II and III.

Paper	Platform	Accession	Reference
I[a]	Mouse Gene 1.0 ST	GEO: GSE54229	Paper I
I[b]	Mouse Exon 1.0 ST	GEO: GSE54228	Paper I
II[c]	Mouse Gene 1.0 ST	AE: E-MTAB-726	Paper II
III[d]	Mouse Exon 1.0 ST	Not Uploaded	

[a] gene expression profiling of primary cell culture model of hypothermia; [b] gene expression profiling of primary cell culture model of hypoxia; [c] gene expression study of hypoxia model on wt and Ngb-null mice; [d] light induced gene expression profiling on wt and Ngb-null mouse retina.

5.2. Microarray data from public domain (Papers I, III)

5.2.1. MicroArray Quality Control dataset (Paper I)

In Paper I we utilized publicly available MicroArray Quality Control (MAQC) project's samples of Human Brain Reference RNA (HBR) by Ambion and Universal Human Reference RNA (UHR) by Stratagen. The data, available on Affymetrix® microarray platforms, was obtained from GEO. Accession numbers and each dataset references are presented in Table 2. Although some datasets contained more than four replicate samples per group, only four robustly chosen technical replicates were used in our analysis.

Table 2. Datasets representing MAQC project's samples

Platform	Accession	Reference
Human Genome U133 Plus 2.0	GEO:GSE9819	(Pradervand et al., 2008)
Human Gene 1.0 ST	GEO:GSE9819	(Pradervand et al., 2008)
Human Exon 1.0 ST	GEO:GSE13069	(Bemmo et al., 2008)

5.2.2. Exposure to light datasets (Paper III)

In Paper III we used two datasets that investigated differential expression upon exposure to a light pulse. Microarray dataset (GEO: GSE29299) had been made available of *ex vivo* retinas obtained from Long Evans rats that were exposed to 3 or 6 hours of light or were left in the dark for the equivalent time (Bedolla and Torre, 2011). In addition, we obtained data (GEO: GSE6904) from a study on

short-term effect of light by measuring differential gene expression from mouse suprachiasmatic nucleus (SCN) after 30 minutes of continuous light pulse or darkness (Porterfield et al., 2007). In *ex vivo* rat retina study the authors utilized Rat Genome 230 2.0 (Affymetrix®) microarrays whereas in mouse SCN study Mouse Genome 430A 2.0 (Affymetrix®) platform was used.

5.2.3. Long-range epigenetic silencing (LRES) data (Paper I)

Coolen et al. (2010) investigated downregulation of neighboring genes over large chromatin regions in cancer cells. For differential gene expression analysis they used gene expression measurements of normal prostate epithelial cells (PrEC) and the prostate cancer cell line (LNCaP). The dataset (GEO: GSE19726) containing two replicate measurements for both conditions were measured with Human Gene 1.0 ST (Affymetrix®) arrays.

5.2.4. Epigenetic modification data (Paper I)

In addition to the above-mentioned LRES gene expression dataset we downloaded additional chromatin immunoprecipitation DNA microarray (ChIP-chip) data (GEO: GSE19726) on epigenetic markers from the same study by Coolen et al. (2010). Histone 3 lysine 9 acetylation (H3K9ac) and Histone 3 lysine 27 tri-methylation (H3K27me3) data in PrEC and LNCaP cells contained MAT scores from two arrays for both cell lines and the epigenetic marker identifier.

6. Reference datasets (Paper I)

6.1. MAQC Taqman® dataset

A dataset of Taqman® assay (GEO: GSE5350) measurements on UHR and HBR samples, performed by MAQC project, were used as a reference to evaluate the performance of differential expression analysis workflows. The normalized expression levels were obtained by the MAQC consortium using the formula $2^{CT_{POLR2A}-CT_i}$ where CT_i refers to the cycle threshold of the gene using the POLR2A gene as reference. Differential expression between UHR and HBR samples was estimated with Student's t-test followed by Bonferroni multiple testing correction.

6.2. HIF-1 and HIF-2 binding sites in hypoxia-induced genes

A list of high-stringency hypoxia inducible factors HIF-1 and HIF-2 binding sites with human genomic coordinates was obtained from the Supplementary Material of Schödel et al. (2011). To cross-link human gene identifiers (RefSeq) with mouse gene identifiers (Ensembl) we downloaded an annotation table from

Ensembl (Cunningham et al., 2014) using R (package biomaRt, (Durinck et al., 2009)). We were able to successfully map 295 mouse orthologs of Hif-1 targets and 245 orthologs of Hif-2 targets. Enrichment of putative Hif-1 and Hif-2 targets among upregulated genes in primary cell culture model of hypoxia was estimated with hypergeometric probability distribution.

7. Microarray data analysis (Papers I–III)

7.1. Microarray preprocessing (Paper I–III)

In Paper I, microarray data preprocessing on MAQC UHR and HBR samples and on data from primary cell culture model was performed using RMA (Irizarry et al., 2003b), FARMS (Hochreiter et al., 2006), DFW (Chen et al., 2007), PLIER (http://media.affymetrix.com/support/technical/technotes/plier_technote.pdf, last access: 30.01.2015) and DEMI. RMA, FARMS and DFW were applied using an R implementation provided in the xps package (Stratowa) downloaded from www.bioconductor.org/packages/release/bioc/html/xps.html (last access: 30.01.2015). PLIER analysis was performed with Affymetrix Power Tools (http://www.affymetrix.com/estore/partners_programs/programs/developer/tools/powertools.affx, last access: 30.01.2015). All methods were applied with default parameters and probe-sets were mapped to genes according to the annotation table downloaded from Ensembl (release 73, (Cunningham et al., 2014)) using biomaRt package in R (Durinck et al., 2009). All other differential gene expression analysis in Papers I, II and III were performed using DEMI.

7.2. Differential expression analysis (Paper I–III)

In addition to DEMI the differential expression analysis on MAQC and primary cell culture model data in Paper I was performed two times on each pre-processing workflow (RMA, FARMS, DFW and PLIER), with Limma (Smyth, 2005) and RankProd (Hong et al., 2006). The analysis was done in R using Limma (http://www.bioconductor.org/packages/release/bioc/html/limma.html) and RankProd (http://www.bioconductor.org/packages/release/bioc/html/RankProd.html) packages in R applied with their default parameters.

If not noted otherwise, differential expression with DEMI was performed with the default settings. However, in Paper III we set more stringent constraints on the differentially expressed genes. More specifically, we set a rule that at least half of the gene-specific probes have to be differentially expressed in the same direction and the minimum cutoff FDR value for the gene had to be below 0.01, instead of the 0.05, which is the default in DEMI.

7.3. Analysis of long-range epigenetic silencing data (Paper I)

A list of putative candidate regions subject to long-range epigenetic silencing (LRES) in prostate cancer was retrieved from Table 1 in the original study (Coolen et al., 2010). Differential expression was estimated with DEMI on 0.5-Mbp genomic windows overlapping with neighboring genomic windows by 50%. Both original comparison and the possible null permutations of the arrays were analyzed. P-values were adjusted for multiple testing using the method published by Benjamini and Yekutieli (2001) and corresponding p-values lower than 0.005 were considered as statistically significant. A DEMI predicted downregulated genomic window was labeled LRES if an overlap of least 0.25-Mbp was observed with an LRES locus from the original publication.

7.4. Estimating overlap of microarray experiments (Paper III)

The overlap of differentially expressed genes representing results from light exposure experiments in Paper III was estimated with Fisher's exact test in R. The p-values were adjusted for multiple testing with FDR (Benjamini and Hochberg, 1995) and enrichment of FDR < 0.05 was considered statistically significant. Since the original data represented measurements from several species (mouse and rat) an annotation table connecting rat gene identifiers to mouse orthologs was obtained from Ensembl (Cunningham et al., 2014) (release 75) using R package biomaRt (Durinck et al., 2009).

8. RNA-seq data and analysis (Paper I)

Two sets of RNA-seq data based on HBR and UHR samples, differing by the number of samples, were obtained from GEO (GSE12946 (Wang et al., 2008) and GSE24283 (Nacu et al., 2011)). RNA-seq reads were aligned with TopHat (Trapnell et al., 2009) (version 1.3.3) to genome indices downloaded from Ensembl (build GRCh37.p12). Differential expression between UHR and HBR samples was estimated using EdgeR (Robinson et al., 2010) (R package, version 3.0.8), DESeq (Anders and Huber, 2010) (R package, version 1.10.1) and Cuffdiff2 (Trapnell et al., 2013) (version 2.0.2). Raw count data, which is required by EdgeR and DESeq was generated with HTSeq (http://www-huber.embl.de/users/anders/HTSeq/, last access: 30.01.2015) (version 0.5.3p9), using flags '-stranded=no -mode=union -type=exon'.

9. Performance analysis (Paper I)

9.1. Performance indicators

We used the following metrics to evaluate the performance of microarray workflows: area under receiver-operator curve (AUC), Matthews correlation coefficient (MCC), true positive rate (TPR), false positive rate (FPR), true negative rate (TNR) and false negative rate (FNR). The latter four were calculated based on the true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) according to the intersection between gene identifiers in the prediction and reference datasets, measured by microarrays and Taqman® assays, respectively. A prediction was classified as TP if the differential expression estimate was significant in both datasets with identical direction (either higher or lower expression in UHR compared to HBR). TN label was assigned to genes with insignificant differential expression in both datasets. Genes that showed significant differential expression in only the reference dataset were termed FN and genes that had significant differential expression in only prediction dataset were labeled as FP.

All the indictors were calculated with R using the ROCR package (http://cran.r-project.org/web/packages/ROCR/, last access: 30.01.2015).

9.2. Microarray simulation test

We created a simulation dataset consisting of 100 simulations of two conditions with four technical replicates, each involving around 45000 genes and around 1.3 million probes that match to transcriptome on Mouse Exon 1.0 ST array. For 1000 randomly chosen genes new signal intensities were simulated by adding a fold change of 2 or –2 to the initial $\log_2$-transformed intensity values for 80% of the target-specific probes. Additional noise was added by adding the same fold-change to a randomly selected 10% of remaining probes. Using DEMI we calculated FDR and MCC by changing two parameters, u (maximum number of probes matching to a target) and t (the maximum number of distinct targets a single probe can match to). In baseline settings DEMI does not force such restrictions.

In a supplementary study we did a power simulation test in order to describe the dependence between the number of probes per target and the target-specific differentially expressed probe with ratios $r_i = \{0.3, 0.5, 0.8\}$. Using the same fold change modification and background noise ratio as in the first simulation test, we calculated the power of Fisher's exact test in correctly identifying differentially expressed targets averaged over 1000 simulations.

9.3. Microarray permutation

To evaluate the performance of microarray analysis methods in conditions where there are few technical replicates available, we performed a permutation analysis using the MAQC samples that contained four technical replicates in both UHR and HBR groups. To that end, all possible combinations of technical replicates with sample size $N = \{3,2\}$ were created. Each permutation of the test sample was compared against all permutations of the reference samples of the same sample size. For every such comparison performance metrics were calculated by averaging across the selected FDR cutoffs (0.05 and 0.01). To evaluate the performance of the entire workflow, at every sample size, the indicators were averaged across all permutations.

10. Functional annotation analysis (Papers I–III)

In Paper II and III functional annotation analysis was performed with g:Profiler (Reimand et al., 2011) using the default settings. In Paper I, DEMI utilized its internal functional enrichment analysis on differentially expressed genes. To that end, the child categories of all GO terms were downloaded using Ensembl Perl API (http://www.ensembl.org/info/docs/api/index.html, last access: 30.01.2015). Ensembl gene identifiers were downloaded using biomaRt (Durinck et al., 2009) package that provides a programmatic access to the Ensembl's (Cunningham et al., 2014) Biomart database. Gene lists corresponding to each GO category was compiled from gene identifiers associated with the category and its children.

RESULTS

1. Paper I

1.1. Evaluation of DEMI performance on simulated expression data

Differential gene expression was simulated by adding a fold-change of 2 or –2 to 80% of on-target probes for 1000 genes selected at random. Equivalent amount of noise was added to a randomly selected 10% of probes remaining on the array. By altering parameters u (the maximum number of probes matching to a target) and t (the maximum number of distinct targets matching to a single probe), we observed that if probes were restricted to match only a single distinct target ($t = 1$), the MCC values were considerably better, as fewer false positives were detected. It indicates that sequence homology between genes can trigger false positive findings if their on-target probes intersect. Limiting the apparent number of on-target probes by setting u to 30 had a smaller but still positive effect on performance by further decreasing the number of false positive findings (Figure 1). Since lowering u reduces sensitivity, it will also reduce the effect of intersecting probes on triggering false positive findings for homologous genes.

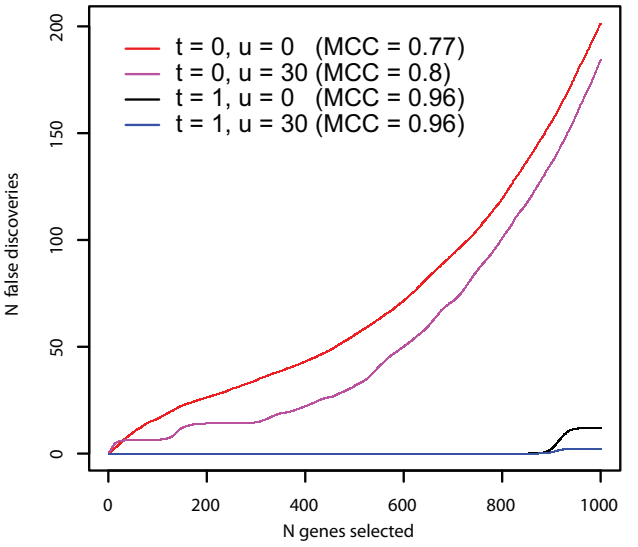

Figure 1. Performance of DEMI on simulated gene-expression data. For each parameter setting the average of 100 simulations are plotted.

DEMI expects that there are several target-specific probes on the microarray. Depending on the target category, this might not always be the case. One possible shortcoming of DEMI is the varying number of probes n_i per target i.

Intuitively, a target with many probes can be assessed more confidently than targets with fewer probes. From the power simulations we observed that the sensitivity to detect differential expression of targets with very few on-target probes ($n \leq 3$) is around 60% even if the enrichment of differentially expressed probes is quite high ($r_i = 0.8$ versus $r_{reference} = 0.1$) (Figure 2). Although potentially a concern, this is not a major problem in real life as such targets form less than 5% of genes (Figure 3). On the other side of the scale are targets with very many on-target probes ($n_i > 40$). In such cases the power to detect small differences in the proportions of differential expression ($r_i = 0.3$ versus $r_{reference} = 0.1$) is around 90% (Figure 2). Very high sensitivity can lead to false positive findings as, for example, $r_i = 0.3$ has substantially more evidence against differential expression of target i even if significantly different from the reference rate. In light of this, we introduced an optional upper limit u for each target i with $n_i > u$. Limiting sensitivity by adjusting all such targets according to the formula $X'_i = X_i * \frac{u}{n_i}$, where X_i is the number of differentially expressed probes per target i, reduces the likelihood of falsely rejecting null hypothesis. Based on our experience, a suitable value for u is ~30, which is also close to the median number of probes per genes on high-density Affymetrix® microarrays. Since the significance of r_i is calculated in relation to the background rate of differentially expressed probes, DEMI is becoming increasingly sensitive when the signal profiles between samples become increasingly similar.

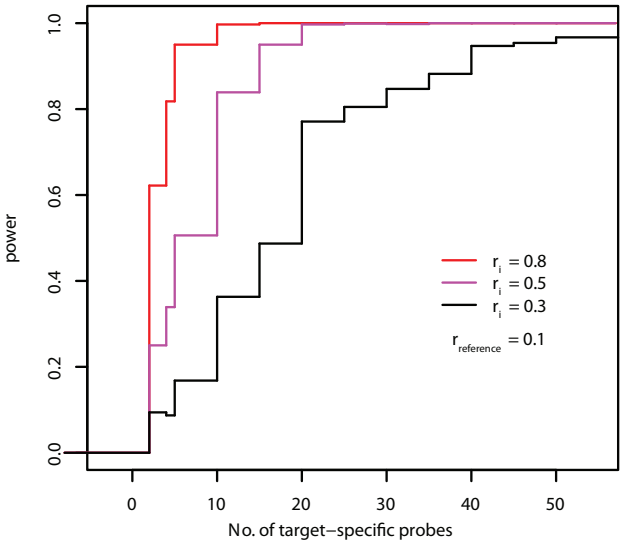

Figure 2. Power analysis of Fisher's exact test. For each combination of input ratios, the average of 1000 simulations is plotted.

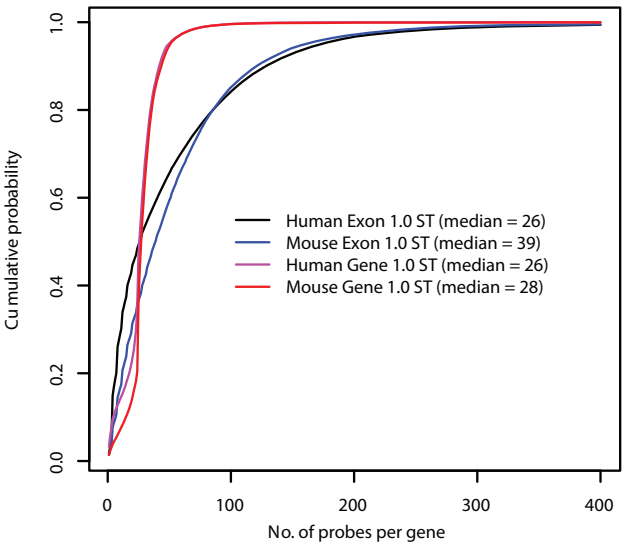

Figure 3. Distribution of gene-specific probe counts with medians presented for four microarrays.

1.2. Comparison of differential expression analysis workflows

Using MAQC UHR and HBR samples, DEMI was compared with eight microarray analysis workflows consisting of four preprocessing strategies (RMA (Irizarry et al., 2003b), FARMS (Hochreiter et al., 2006), DFW (Chen et al., 2007) and PLIER (Affymetrix, 2005)) combined with two differential expression estimation methods (Limma (Smyth, 2005) and RankProd (Hong et al., 2006)). In addition, RNA-seq data on the same samples was analyzed with EdgeR (Robinson et al., 2010), DESeq (Anders and Huber, 2010) and Cuffdiff 2.0 (Trapnell et al., 2013). Each workflow was assessed by calculating several performance metrics (MCC, AUC, TPR, FPR, TNR and FNR) with varying size of the input dataset ($N = \{2, 3, 4\}$). Performance was estimated in relation to the "gold-standard" based on results from Taqman® assays targeting 867 genes as published by the MAQC consortium (Consortium et al., 2006). Each metric was summarized as an average over two FDR cutoffs (0.01, 0.05) across all permutations of the specified sample size. All performance indicators were plotted juxtaposed on a radial performance plot (Figure 4). A large colored plane on the upper semicircle of the plot is characteristic of good performance, as indicated by high MCC, AUC, TPR and TNR values, while coloring in the lower semicircle demonstrates bad performance. Ideally, the colored plane would occupy most of the upper semicircle while leaving the lower semicircle blank.

DEMI demonstrated most stable performance across different sample sizes and microarray platforms. RankProd had a high true negative rate but it came at the cost of a high false negative rate. Thus, it appears that RankProd

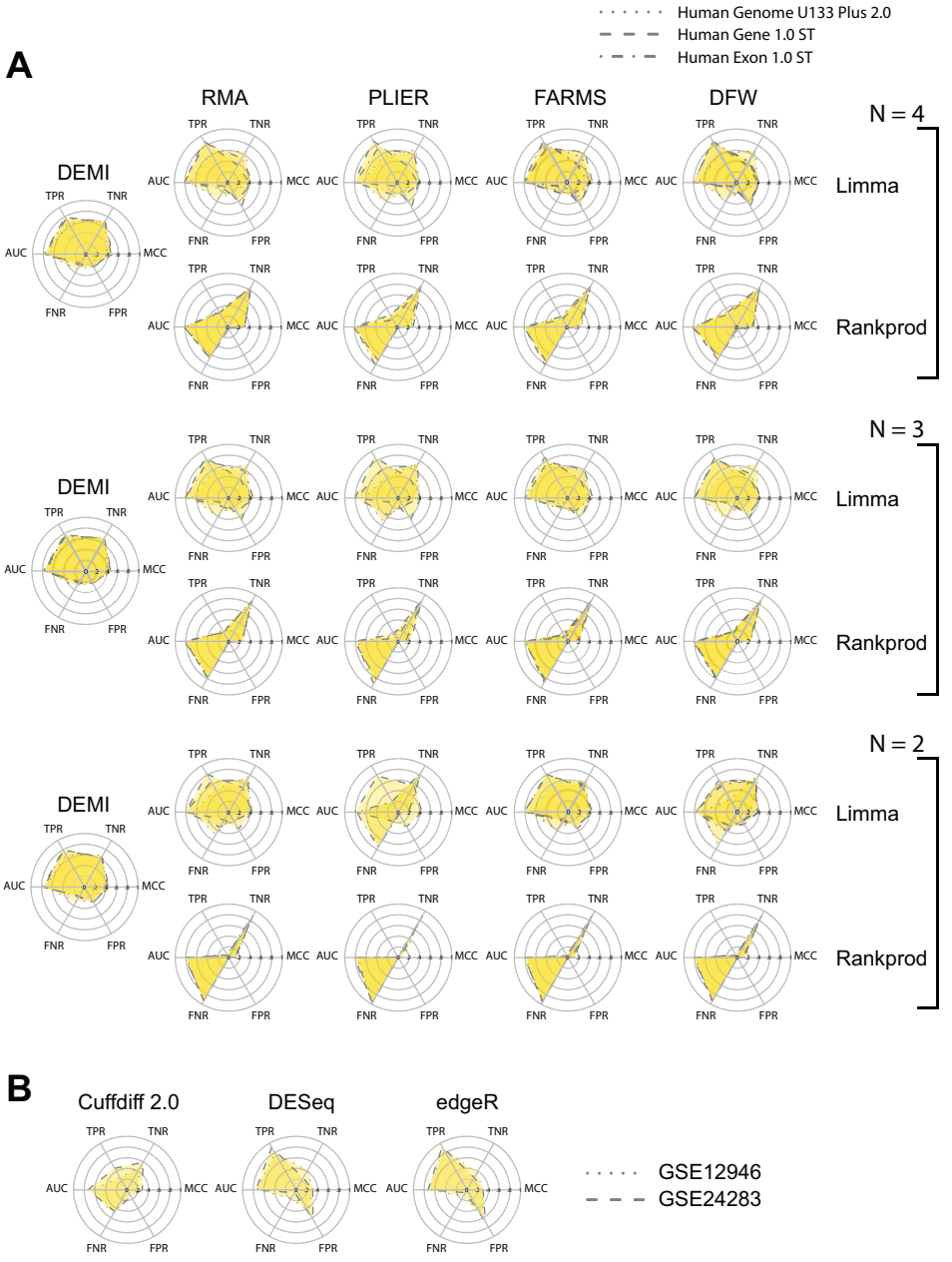

Figure 4. Performance analysis of differential expression analysis workflows on MAQC projects samples. Performance of each pipeline is presented as a radial plot, which includes six complementary performance indicators. Results from three microarray platforms (**A**) and two RNA-seq datasets (**B**) are presented. Abbreviations: AUC – Area Under Curve, MCC – Matthews Correlation Coefficient, TPR – True Positive Rate, TNR – True Negative Rate, FPR – False Positive Rate and FNR – False Negative Rate.

was lacking in power at the given FDR cutoffs, which led it to produce more negative findings than was the case for DEMI and Limma. In contrast, Limma had the highest TPR at sample sizes $N = \{4, 3\}$ and similar MCC values to DEMI. From the performance plot (Figure 4A) it appears that PLIER is less consistent across different array types than other normalization methods, whereas the performance of DEMI is rather stable regardless of the microarray platform or the sample size.

In comparison to top-performing microarray analysis methods, RNA-seq analysis workflows showed somewhat lower MCC values (Figure 4B). Although DESeq and EdgeR were more sensitive at detecting differentially expressed genes, more false positives genes were detected. This resulted in lower MCC values (~0.3) when compared to top-performing microarray analysis workflows. Cuffdiff 2 was the most conservative out of the three RNA-seq differential analysis methods with substantially lower TPR and a very low FPR.

1.3. Performance of DEMI in cell culture model of hypoxia

DEMI was compared with eight other microarray analysis workflows in accurately detecting differentially expressed genes in cell culture model of hypoxia. We reasoned that real-life experimental data should be a more valid basis for benchmarking than the somewhat artificially constructed MAQC reference RNA samples. To this end, we conducted an experiment where differential expression was estimated in response to 1% oxygen for 24 hours in comparison to atmospheric oxygen levels in the cell culture. The transcriptional mechanisms of hypoxia response are extensively studied and well characterized (Greer et al., 2012; Semenza, 2012), allowing us to validate our cell culture model and the accuracy of microarray analysis workflows. Specifically, we searched for the upregulation of Gene Ontology (GO) (Ashburner et al., 2000) categories ‘cellular response to hypoxia’ (GO:0071456) and ‘glycolysis’ (GO:0006096) among the differentially expressed genes in response to prolonged hypoxia (Lendahl et al., 2009). When three or more replicate samples were available almost all microarray analysis workflows showed consistency in producing an enrichment of the aforementioned GO categories among up-regulated genes (Figure 5). However, only DEMI was able to correctly detect the enrichment of hypoxia response pathways in all possible subsets when technical replicate sample size was set to 2.

In addition, we calculated the enrichment of differentially expressed genes among mouse orthologs with HIF-1 and HIF-2 binding sites. Similar to pathway enrichment analysis, most methods showed significant enrichment for Hif-1 and Hif-2 targeted genes among significantly upregulated genes when replicate sample size was $N \geq 3$ (Table 3). Once again, only DEMI produced a list of differentially expressed genes yielding significant enrichment of Hif-1 and Hif-2 target genes when sample size was restricted to only two technical replicates.

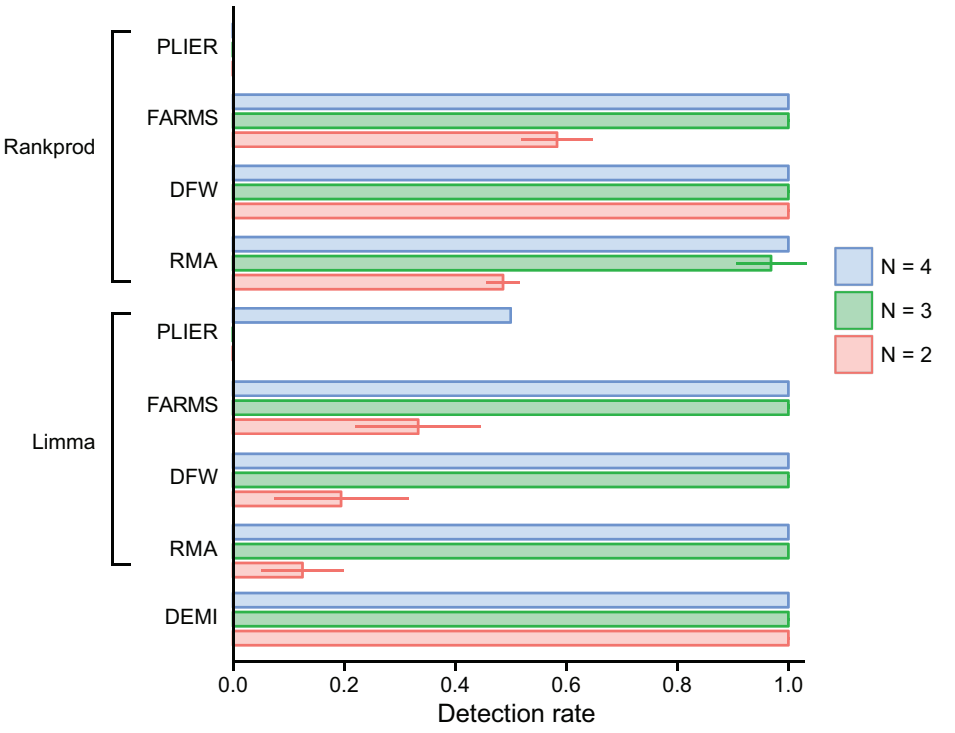

Figure 5. Performance analysis of differential expression analysis workflows on data from mouse embryonic fibroblasts exposed to 1% O2 for 24 h. The plot depicts the combined detection rate of GO categories 'cellular response to hypoxia' (GO:0071456) and 'glycolysis' (GO:0006096) as indicators of hypoxia response. The data is plotted as mean ± standard error of all possible comparisons between subsets of size N of the hypoxic and normoxic groups (original N = 4).

Table 3. Enrichment of mouse orthologs of HIF-1 targets among significantly up-regulated genes in mouse embryonic fibroblasts exposed to 1% O_2 for 24h.

Normali-zation	DE	N = 4		N = 3		N = 2	
		Mean	SEM	Mean	SEM	Mean	SEM
Relative ranking	DEMI	3.0E-25	NA	4.3E-22	2.4E-22	6.0E-19	5.4E-19
DFW	Limma	1.4E-28	NA	3.5E-23	2.2E-23	0.617	0.081
DFW	RankProd	1.4E-30	NA	2.2E-23	1.5E-23	8.2E-08	4.8E-08
FARMS	Limma	9.5E-22	NA	2.5E-21	1.9E-21	0.457	0.082
FARMS	RankProd	6.8E-20	NA	3.1E-10	2.1E-10	0.037	0.011
PLIER	Limma	9.4E-10	NA	0.9	0.1	1	0
PLIER	RankProd	3.5E-01	NA	1	0	1	0
RMA	Limma	6.0E-19	NA	1.7E-14	1.5E-14	0.481	0.083
RMA	RankProd	1.4E-18	NA	4.3E-13	2.1E-13	0.001	4.6E-04

Differential expression was estimated with nine pipelines including various preprocessing and differential expression analysis methods. The data is presented as mean and standard error of hyper-geometric p-values from all possible comparisons between subsets of size N of the hypoxic and normoxic groups (original N=4). N, sample size; SEM, standard error of mean; NA, not available.

1.4. Evaluation of gene expression dynamics in response to hypothermia

To demonstrate that DEMI can handle unconventional experimental designs, we investigated differential gene expression in response to increasing duration of mild hypothermia (32°C). Hypothermia is a clinically proven treatment for various hypoxic and ischemic conditions (Bernard et al., 2002; Miyazawa et al., 2003; Polderman, 2008). We reasoned that a time-dependent near-monotonic increase or decrease of gene expression is a good indicator of a possible causal relationship between the treatment and transcriptional regulation. In addition, this experimental design required only a single microarray for each condition at each time point, which is much less than the suggested minimal number of replicates (reviewed in Allison et al., 2006) per treatment for factorial designs. Using Kendall's tau statistic, a measure of rank correlation, DEMI identified 1750 and 274 genes exhibiting a significant monotonic-like temporal response either to hypothermia or normothermia, respectively. Among the top five genes with upregulated expression under hypothermic conditions, we found Cold inducible RNA-binding protein (*Cirbp*) (Figure 6), which is a well-known hypothermia-responsive gene (Fujita, 1999; Nishiyama et al., 1997). Of interest, we also identified several genes related to the antioxidant systems, which appear to be induced under hypothermia (Table 4).

Table 4. Genes responding to increasing duration of hypothermia with significantly monotonic increase in expression and relating to the antioxidant system.

Gene ID	Symbol	P-value	FDR	System
ENSMUSG00000003849	Nqo1	9.04E-35	4.72E-30	Quinone detoxification
ENSMUSG00000032802	Srxn1	1.23E-19	2.14E-16	Glutathione
ENSMUSG00000027610	Gss	1.08E-13	4.74E-11	Glutathione
ENSMUSG00000020250	Txnrd1	4.29E-11	9.83E-09	Thioredoxin
ENSMUSG00000032350	Gclc	4.29E-11	9.83E-09	Glutathione
ENSMUSG00000000811	Txnrd3	9.85E-09	1.19E-06	Thioredoxin

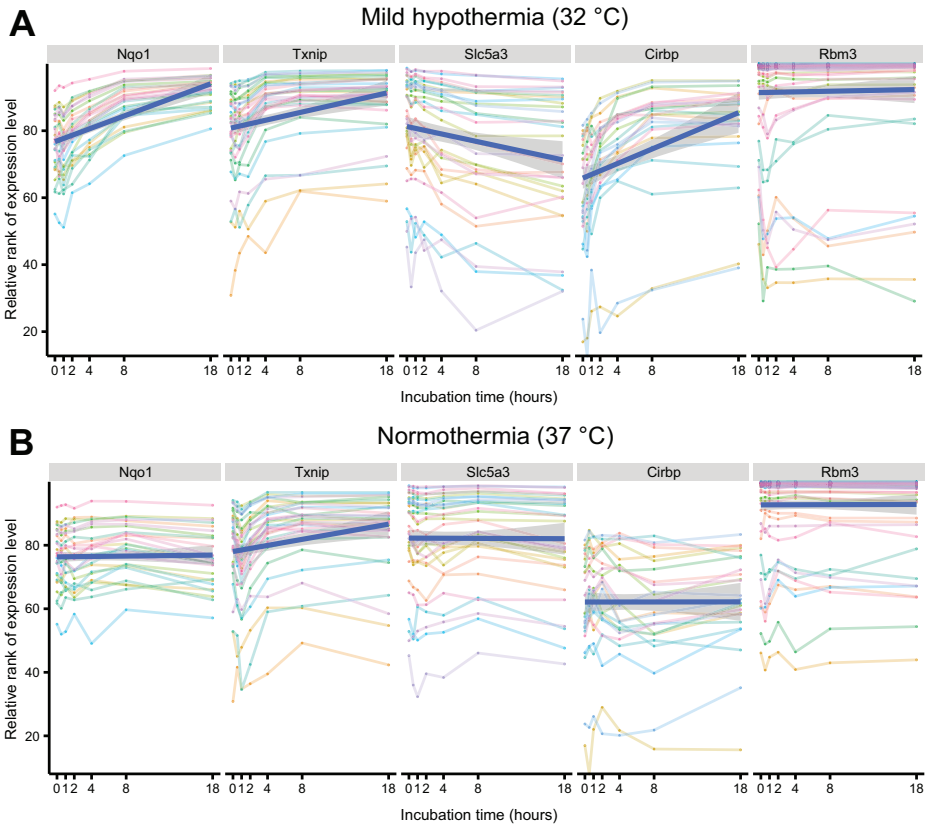

Figure 6. Large-scale analysis of temporal dynamics of transcription in mouse embryonic fibroblasts exposed to mild hypothermia (**A** and **B**). Temporal profiles of probe expression levels of selected genes during mild hypothermia (**A**) and normothermia (**B**). The solid blue line indicates a linear fit to the data points and the gray shadowing represents standard error of the fit.

1.5. Studying long-range epigenetic silencing (LRES) in cancer

In order to demonstrate that DEMI can be used to detect differential expression of unconventional target categories, we sought to identify genomic regions that were subject to long-range epigenetic silencing (LRES) in the publicly available microarray data published by Coolen et al. (2010). The dataset consisted of four microarrays, two replicates for prostate cancer cell line (LNCaP) and two replicates for normal prostate epithelial cells (PrEC). Under these conditions DEMI was able to identify 2242 genomic regions (each spanning 0.5 Mbp) as downregulated in prostate cancer from a total of 22,630 partially overlapping regions. In array permutations corresponding to the null hypothesis (arrays of normal cells and cancerous cells were mixed), only 222 and 201 genomic regions were found to be significantly downregulated. To validate the correct identification of LRES regions, we calculated the enrichment of putative LRES regions reported in the original study (Coolen et al., 2010) within downregulated genomic regions predicted by DEMI. Out of 47 putative LRES regions 38 had been identified by DEMI, yielding an FDR adjusted p-value of $3.04e\text{-}13$ (Fisher's Exact Test). In contrast, in permutations corresponding to null-hypothesis, adjusted p-values 0.961 and 1 were observed. In addition, we examined whether the downregulated regions predicted by DEMI were enriched for altered chromatin modifications using ChIP-chip data from the original study of Coolen et al. (2010). More specifically we analyzed the levels of histone 3 lysine 27 tri-methylation (H3K27me3) and histone 3 lysine 9 acetylation (H3K9ac) on genomic regions inspected by DEMI. Increased levels of H3K27me3 are known to be more prevalent at silent promoters (Barski et al., 2007) whereas acetylation has been associated with transcriptionally active chromatin (Rice and Allis, 2001). In the original paper the authors witnessed that the loss of H3K9 acetylation led to a reduction in gene transcription (Coolen et al., 2010). This finding was also confirmed by our analysis where downregulation of H3K9ac was associated with silenced genomic regions (p = 5.4e-22, hypergeometric probability distribution) and upregulation of H3K9ac corresponded to upregulated DEMI regions (p = 7.99e-05, hypergeometric probability distribution). We did not observe significant association between downregulated genomic regions and H3K27me3.

2. Paper II

2.1. Studying transcriptional response to hypoxia with DEMI

The study was undertaken to identify possible effects of Ngb deficiency on hypoxia response in the mouse brain. Mice were subjected to acute (90 min) or prolonged (24h) hypoxia (7% oxygen) followed by transcriptome-wide analysis of differential gene expression in the whole brain.

Transcriptional response to hypoxia was investigated in both genotypes individually. At 90 minutes of hypoxia 517 upregulated and 347 downregulated transcripts were identified in wt mice, whereas the numbers in Ngb-null mice were 404 and 364 respectively. Under prolonged hypoxia of 24 hours the numbers of differentially expressed genes increased, 997 upregulated and 1862 downregulated in wt mice and 1535 upregulated and 1573 downregulated in Ngb-null mice. Some of the genes significantly upregulated in both genotypes are previously known to be hypoxia responsive (Table 5, Figure 8 A-B). Hypoxia-inducible factor 1-alpha (*Hif1A*) was significantly upregulated only in Ngb-null mice after 24 hours of hypoxia when compared to naive Ngb-null mice ($p < 0.001$, respectively).

Table 5. Previously known hypoxia responsive genes that were detected as upregulated by hypoxia

Gene name	Duration	Reference
Ier3	90 min	(Blais et al., 2004)
Bhlhe40	90 min	(Yun et al., 2002)
Cdkn1A	24 h	(Denko et al., 2000)
Bnip3	24h	(Kothari et al., 2003)
Mt2	24 h	(Murphy et al., 1994)
Vegfa	24 h	(Lu and Kang, 2010)
Slc2a1	24 h	(Ebert et al., 1995)
Kdm3a	24 h	(Wellmann et al., 2008)

Using qPCR, we confirmed hypoxia dependent upregulation of the following genes in both genotypes (Figure 7): *Ier3* (90 min, 24 h), *Hif1A* (90 min), *Mt2* (24 h) and *Cdkn1A* (24 h). *Chd7* and *Cygb* were shown to be exclusively up-regulated only in Ngb-null mice at both 90 minutes and 24 hours. The overlap between the microarray and qPCR results is an indication of DEMI's accuracy. Inter-genotype agreement of transcriptional response was also observed with functional annotation analysis of hypoxia responsive genes in relation to KEGG pathways (Kanehisa and Goto, 2000) such as "apoptosis" (Hedtjarn et al., 2004), "oxidative phosphorylation" (Rodriguez-Enriquez et al., 2010), "mTOR signaling" (DeYoung et al., 2008) and "VEGF signaling pathway".

To identify transcripts with coherent evidence of differential expression we calculated the product of differential expression p-values from both wt and Ngb-null genotypes and ranked-ordered them for each time-point separately (Figure 8 C-D). The five highest-ranking genes after 90 minutes of hypoxia

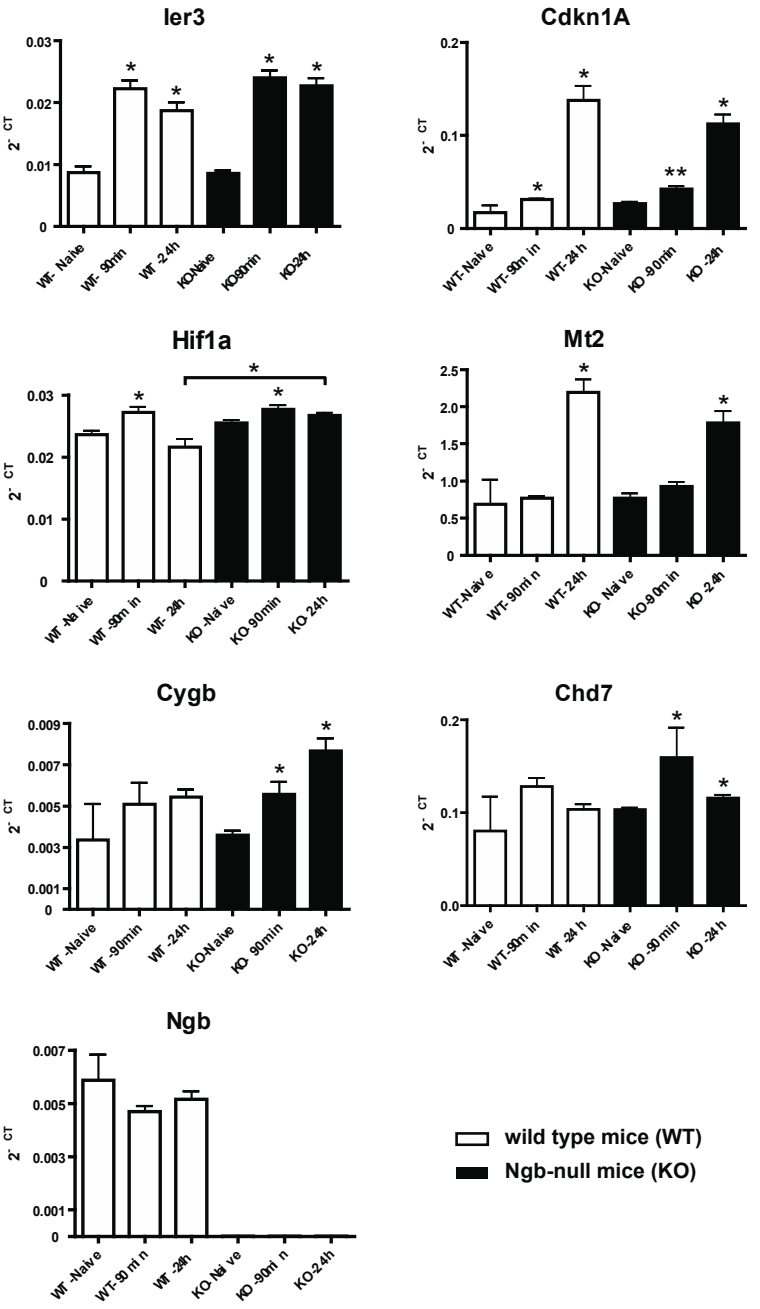

Figure 7. qPCR measurements of differentially expressed genes as detected by microarray analysis. Data is plotted as mean +/– standard error. Asterisk denotes a statistically significant difference in gene expression between the group bearing the asterisk and the naive group of the same genotype. *p<0.05, **p<0.001 (Mann-Whitney test). Genotypes: wild-type (white-bars), Ngb-null (black bars).

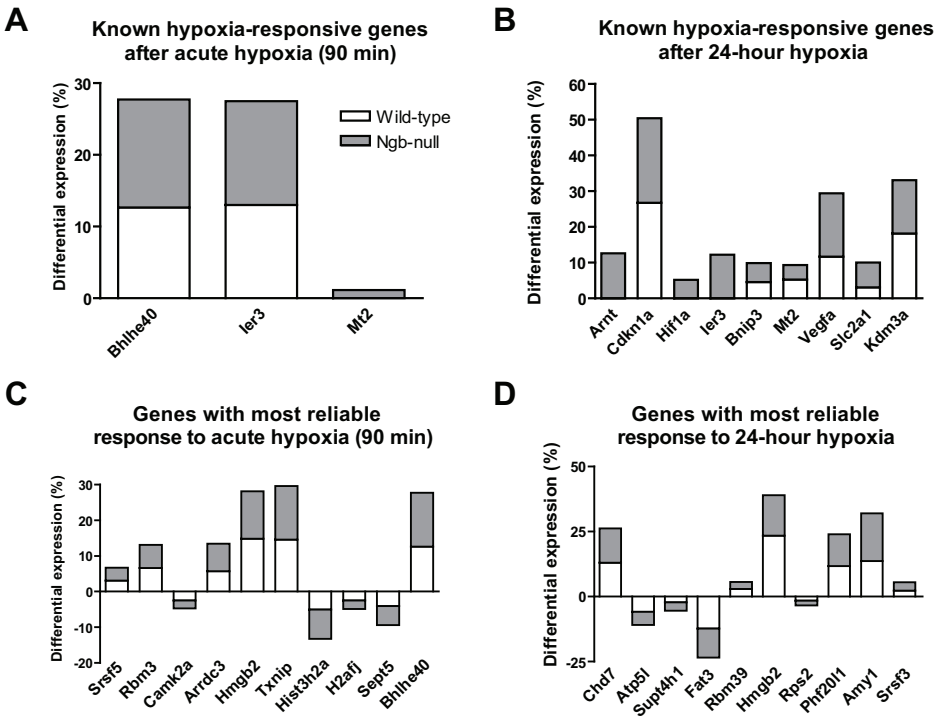

Figure 8. Differential gene expression in response to hypoxia. **A-B** Differential expression ($p<0.05$) of several well-known hypoxia responsive genes after acute (**A**) and 24 hours of hypoxia (**B**) according to Affymetrix Mouse Gene 1.0 ST array. Bars represent the degree of differential expression between hypoxic and normoxic mice of the respective genotype based on the difference in the mean relative rank of target-specific probes (wt mice – white bars, Ngb-null mice – gray bars). **C-D**: Genes with most reliable differential expression in response to acute (**C)** and 24 hour (**D)** hypoxia based on the product of differential expression p-values in both genotypes. Genes are ordered from left to right by the ascending product of differential expression p-values (i.e. decreasing reliability of the differential expression estimate). Bars represent differential expression between hypoxic and normoxic mice of the respective genotype based on the difference in the mean relative rank of the differentially expressed probes (wild type mice – white bars, Ngb-null mice – gray bars).

were *Srsf5* (serine/arginine-rich splicing factor 5), *Rbm3* (RNA binding motif protein 3), *Camk2a* (calcium/calmoduline-dependent protein kinase II alpha), *Arrdc3* (arrestin domain containing 3) and *Hmgb2* (high mobility group box 2). At 24 hours of hypoxia the five top-ranging genes were *Chd7* (chromodomain helicase DNA binding protein 7), *Atp5l* (ATP synthase, H+ transporting, mitochondrial F0 complex, subunit g), *Supt4h1* (suppressor of Ty 4 homolog 1), *Fat3* (FAT tumor suppressor homolog 3) and *Rbm39* (RNA binding motif protein 39).

2.2. Effect of Ngb deficiency in response to hypoxia

Only 22 transcripts were identified as differentially expressed between untreated wt and Ngb-null mice. More compellingly, they were all downregulated in Ngb-null mice. To evaluate genotype-dependent transcriptional response on the onset of hypoxia we searched for genes with consistent differential expression at hypoxic time-points (90 min and 24 h) but not in naive conditions (t0). We observed an upregulation of 181 genes and downregulation of 121 genes in Ngb-null mice. Three most significant upregulated genes ($p < 0.001$) in Ngb-null mice during hypoxia were *Cspp1* (centrosome and spindle pole associated protein 1), *Adi1* (acireductone dioxygenase 1) and *Prpf4b* (PRP4 pre-mRNA processing factor 4 homolog). The three most significantly downregulated genes ($p<0.001$) in Ngb-null mice under hypoxia were *Ubc* (ubiquitin C), *Rplp0* (ribosomal protein, large, P0) and *Kidins220* (kinase D-interacting substrate 220).

To identify genotypic differences in gene expression with respect to the duration of hypoxia we conducted differential expression analysis of wt and Ngb-null mice independently at either 90 minutes or 24 hours of hypoxia. Only 6 transcripts were observed to be upregulated in Ngb-null mice after 90 minutes of hypoxia, whereas the number of downregulated transcripts was 344. The six upregulated transcripts represented *Chd7*, *Snap23*, *Llph* and *Hnrnpa1* genes. Functional annotation analysis revealed downregulation of pathways in Ngb-null mice that were related to signal transduction ("IkB is ubiquitinated and degraded") and metabolism ("glyceraldehyde-3-phosphate dehydrogenase activity"). After 24 hours of hypoxia the numbers were reversed with 204 genes upregulated and 13 genes downregulated in Ngb-null mice compared to wt. No functional annotation was observed for these genes.

3. Paper III

3.1. Studying transcriptional response in Ngb-deficient mouse retina with DEMI

The study was undertaken to identify possible effects of Ngb deficiency on light-induced gene expression response in the mouse retina. The retinas were either dark-adapted (t0) or exposed to light for 1.5 or 5 hours. To identify robust genotype-dependent differences, arrays over all time-points were pooled ($n = 15$) and analyzed for differential expression using DEMI. Several genes were found to be differentially expressed (Table 6), of which *Akap6*, *Entpd5* and *Atp8a2* were confirmed with qPCR (Figure 9).

It has been recently shown that Ngb can affect the expression levels of *Hif1α*, *Nrf2* (*Nfe2l2*) and antioxidant response in hypoxia conditioned cells (Hota et al., 2012). Therefore we chose to examine the expression levels of the following genes in dark-adapted retina with qPCR: *Hif1α*, the *Hif1α* target gene *Bnip3*

(Mendez et al., 2010), the glutathione system genes *Gsr* and *Gstz1*, the master regulator of oxidative stress responsive genes *Nfe2l2* (Linker et al., 2011) and the unfolded protein response regulator *Xbp1* (Iwakoshi et al., 2003). None of the genes were differentially expressed between wt and Ngb-null genotypes in dark-adapted retina (Figure 10), indicating that there was no apparent increase of ER-stress, oxidative stress or hypoxia response in the Ngb-null retina.

Table 6. Five most significantly up- and downregulated genes between wt and Ngb-null mouse retina when light exposure was ignored.

	FDR	Symbol	Description
↑	1.10e-23	*Akap6*	A kinase (PRKA) anchor protein 6
↑	6.76e-22	*Tmem229b*	Transmembrane protein 229B
↑	2.60e-22	*Serpina3n*	Serine (or cysteine) peptidase inhibitor, clade A, member 3N
↑	1.28e-13	*Gdpd3*	Glycerophosphodiester phosphodiesterase domain containing 3
↑	1.28e-13	*Ccdc115*	Coiled-coil domain containing 115
↓	7.56e-36	*Entpd4*	Ectonucleoside triphosphate diphosphohydrolase 4
↓	2.39e-31	*Atp8a2*	ATPase, aminophospholipid transporter-like, class I, type 8A, member 2
↓	5.10e-18	*Snapc1*	Small nuclear RNA activating complex, polypeptide 1
↓	2.20e-16	*Heatr5a*	HEAT repeat containing 5A
↓	2.37e-13	*Lcmt2*	Leucine carboxyl methyltransferase 2

↑ gene expression was higher in Ngb-null mice; ↓ gene expression was lower in Ngb-null mice; FDR – false discovery rate.

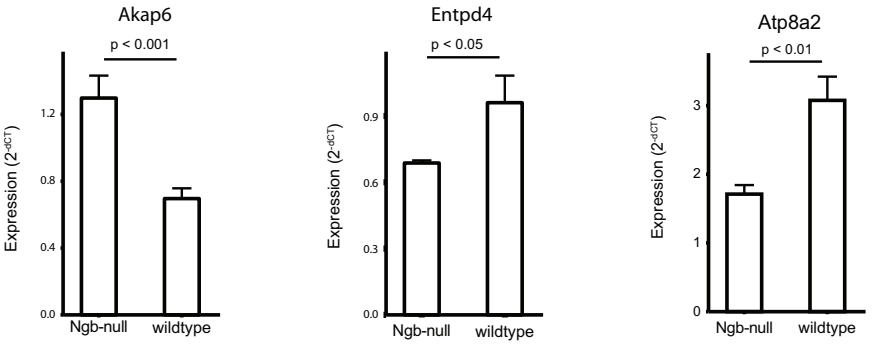

Figure 9. qPCR verification of differentially expressed genes between Ngb-null and wt mouse retina. Expression was normalized to Hypoxanthine phosphoribosyltransferase 1 (Hprt1) as the internal reference (n=9–10). Differential expression was estimated with t-test to confirm the direction of expression dynamics reported by microarray.

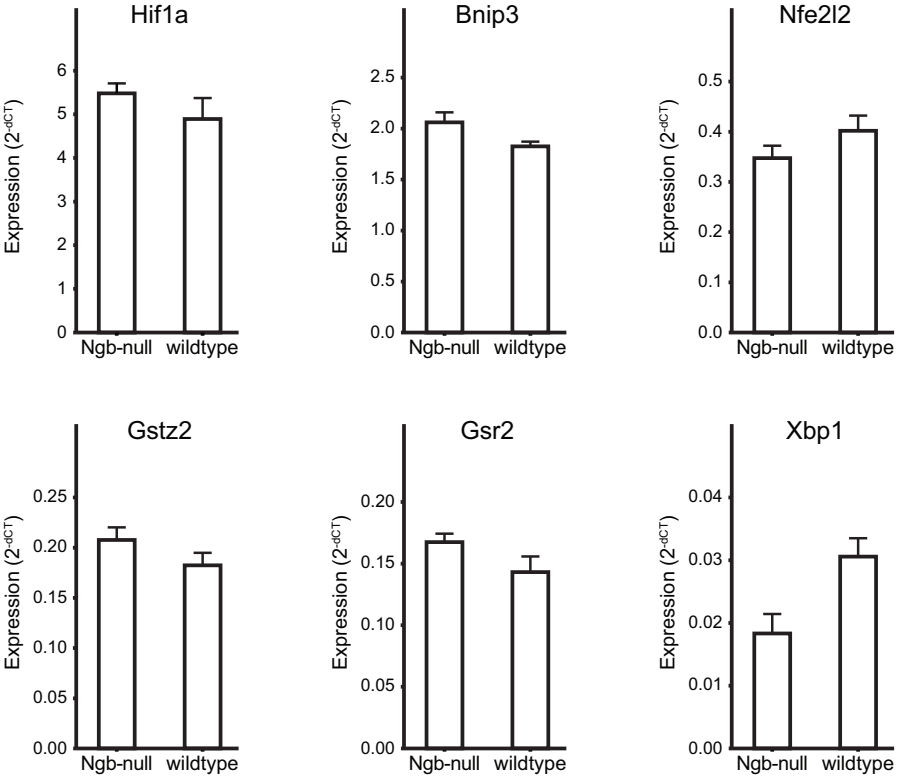

Figure 10. qPCR measurements of hypoxia (Hif1A, Bnip3), oxidative stress (Nfe2l2, Gstz2, Gsr2) and unfolded protein response (Xbp1) marker genes in dark-adapted retina. No differential expression was detected between the genotypes (t-test, n=4–5). Expression was normalized to Hypoxanthine phosphoribosyltransferase 1 (*Htpr1*) as the internal reference.

3.2. Studying transcriptional response to light with DEMI

Compared to the number of differentially expressed genes between genotypes, prolonged light exposure had a considerably stronger effect on the numbers of differentially expressed genes (Table 7).

Results from our study were compared with two other publicly available datasets that studied the effect of light on gene expression. In the first study, *ex vivo* retinas from adult Long Evans rats exposed to 3 or 6 hours of light were compared against equivalent dark-conditioned retinas (Bedolla and Torre, 2011). The second dataset consisted of gene expression measurements from mouse suprachiasmatic nucleus (SCN) taken after 30 minutes of light and sham-light pulse (Porterfield et al., 2007). An overlap between the differentially expressed genes from different datasets was evaluated using hypergeometric probability distribution. Statistically significant overlap was observed between

the results from our study and external datasets (Table 8) validating our experimental results. Specifically, a significant overlap between the upregulated genes in the mouse retina after 1.5 h light exposure and in the mouse SCN after 0.5 h light pulse was observed (Table 8). Similarly, a significant overlap of upregulated genes after 5 h of light exposure in mouse retina with data from 6 h light pulse (Table 8) in *ex vivo* rat retina cross-validated our analysis. In contrast, no significant overlap of downregulated genes was observed between the datasets, indicating that light-specific response is mostly associated with the induction of gene expression. As a negative control, no significant overlap between genes differentially expressed in opposite directions was observed. Enrichment of light-inducible genes from external datasets among light induced genes in the current study was similar in wt and Ngb-null mice. This suggests that the transcriptional response to light is largely unaltered in Ngb-deficient retina.

Table 7. Number of differentially expressed genes in the mouse retina (FDR < 0.01)

Light pulse	**Genotype differences**[a]		**Ngb-null mice**[b]		**Wildtype mice**[b]	
0 h	7	6	–	–	–	–
1.5 h	23	27	5	36	28	11
5 h	9	11	21	29	323	283
	Up	**Down**	**Up**	**Down**	**Up**	**Down**

[a] number of differentially expressed genes between Ngb-null and wt mice after different durations of light exposure; [b] differential expression between light-adapted (1.5 h or 5 h) and dark-adapted retina (0 h).

When genotype was ignored and arrays from wt and Ngb-null mouse retina of corresponding time points were pooled, we identified several early response genes like *Egr1*, *Fos*, *Fosl2*, *Per1* and *Nr4a1-3*, to be upregulated after 1.5 hours of light exposure (Figure 11). Of these *Nr4a1*, *Per1*, *Fos*, *Fosl2* and *Nr4a2* have also been confirmed in the study by Araki et al (Araki et al., 2006) as light inducible transcripts in SCN, which is a light-responsive brain area directly innervated by the eye. In addition, by using DEMI for the reanalysis of the data from Porterfield et al. (2007) who studied light response in SCN, we observed an upregulation of *Fos*, *Nr4a1* and *Egr1* after 0.5 hour light exposure (Figure 11). Functional annotation analysis of upregulated genes at 5 hours of light pulse revealed significantly enriched gene ontology categories related to translation, RNA-splicing and visual perception.

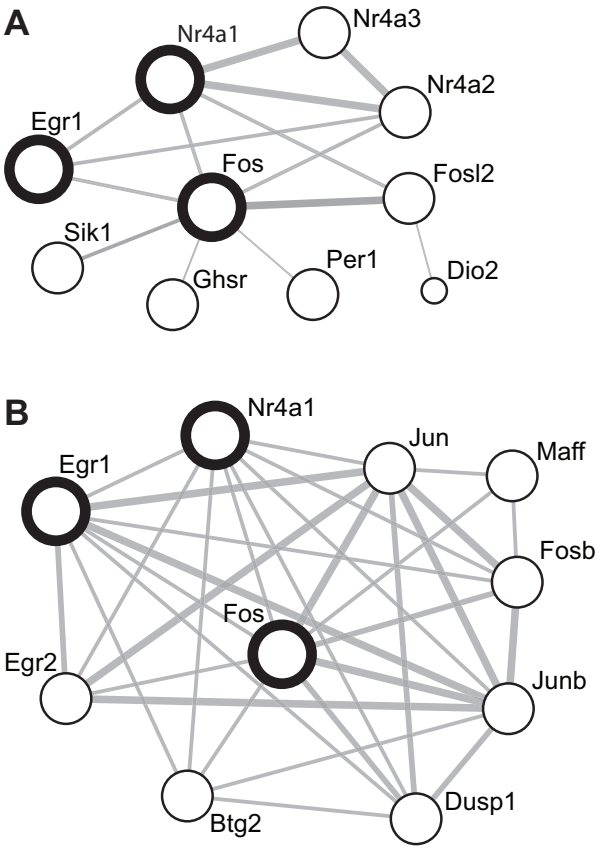

Figure 11. Known and predicted protein interactions between light-induced genes in the retina and the SCN. The circles depict genes, which were upregulated in response to light. Connecting lines indicate protein interaction information as retrieved from STRING (Franceschini et al., 2013). **A.** genes induced by 1.5h light pulse (Ngb-null and wild-type arrays were pooled). **B.** genes induced in the mouse SCN by 0.5 light pulse (reanalysis of data published by Porterfield et al. (2007)). Thicker lines show more and thinner lines show less evidence of interactions between proteins. Genes overlapping between two datasets are presented with accentuated circles.

Table 8. Similarity of gene expression responses in the present study and external datasets

	Light pulse, h	Diff. expression	**Present study: wild-type and Ngb-null pooled**				**Present study: Ngb-null**				**Present study: wild-type**			
			1.5 h		5 h		1.5 h		5 h		1.5 h		5 h	
			up	down	up	down	up	down	up	down	up	down	up	down
Mouse SCN (Porterfield et al., 2007)	0.5	up	1.84e-04	1	0.7	1	1.14e-04	1	1	1	2.02e-05	1	1	1
		down	1	1	0.49	1	1	1	1	1	1	1	0.85	1
Rat retina (Bedolla and Torre, 2011)	3	up	1	1	0.49	1	1	1	0.18	1	1	1	0.17	1
		down	1	0.77	1	0.49	1	1	1	1	1	1	1	0.85
	6	up	0.83	1	7.42e-07	1	0.18	1	1.14e-04	1	0.85	1	2.02e-05	1
		down	1	1	1	0.82	1	1	1	0.68	1	1	1	1

The table lists FDR-values for the enrichment of differentially expressed genes (hypergeometric probability distribution) between the present study and related external data. In order to reveal genotype-dependent effects in the present study, differential expression in light-exposed vs. dark-adapted retina was estimated in both genotypes separately as well as in pooled genotypes. The grey cells indicate that the direction of differential expression in the two datasets is the same. Ngb, neuroglobin; Ngb-null, Ngb-deficient; FDR, false discovery rate.

DISCUSSION

1. DEMI – a novel method for microarray differential expression analysis

This dissertation introduces a novel method called Differential Expression from Multiple Indicators (DEMI). DEMI is a novel probe-level differential gene expression analysis framework that can utilize a variety of statistical tests as "plugins" to evaluate differential expression. It has been applied to different experimental designs including case and control comparisons (Papers II and III) along with time-series design (Paper I). In addition, we have used DEMI to evaluate differential expression on genes, transcripts and genomic regions (Papers I–III). Although several well-established microarray analysis methods can be applied with various experimental designs, DEMI takes a unique approach to the differential expression estimation procedure. Traditionally, probe-level signal intensities are summarized before differential expression testing. In DEMI, however, differential expression is evaluated on the probe-level data. These estimates are then used to evaluate differential expression on the target level. Essentially, DEMI treats probes as voters and, for each target, compares the ratio of target-specific differentially expressed probes to the ratio of differentially expressed off-target probes. As we have shown, it can partially compensate for the loss of power, which occurs at very small sample sizes where statistical estimates become very noisy. Thus, DEMI takes advantage of repeated measurements per target as provided by high-density microarrays to enable more accurate detection of differential expression with few samples.

1.1. Benchmarking of differential expression analysis methods

DEMI was evaluated using several independent datasets with varying experimental designs to demonstrate its applicability in different situations. First, it was benchmarked in relation to eight other microarray differential expression analysis pipelines and three RNA-seq analysis pipelines, using data from Microarray Quality Control (MAQC) project (Consortium et al., 2006). So far, MAQC, an initiative of US Food and Drug Administration (FDA), represents the most extensive effort to characterize the performance of various gene expression measurement platforms and data analysis methods. MAQC consortium used two standardized RNA pools, Universal Human Reference RNA (UHR) and Human Brain Reference RNA (HBR) to generate gene expression data. HBR is high quality RNA extracted from multiple donors and several brain regions to obtain an unbiased and reproducible coverage of the brain transcriptome (https://www.lifetechnologies.com/order/catalog/product/AM6050, last access: 30.01.2015). UHR was obtained by pooling RNA from ten different cell lines (http://www.genomics.agilent.com/article.jsp?pageId=1452, last access: 30.01.2015). Earlier efforts have used spike-in datasets with a Latin Square design (Chen et al., 2007; Hochreiter et al., 2006; Irizarry et al., 2003b),

which have been criticized for the small number of spike-in sequences (typically a couple of dozen) used to make transcriptome-level conclusions about performance (Draghici et al., 2006). Another option is to use dilution studies where known samples are mixed or diluted in known ratios (Consortium et al., 2006). Latin Square datasets are created under controlled conditions providing relatively little variation between experimental and technical conditions and between samples (Gyorffy et al., 2009). On the other hand, when using UHR and HBR for benchmarking, you are dealing with an experiment with large inter-group differences (two highly different pools) and very small intra-group differences, represented by technical replicates without biological variation. Such artificial data might not be representative enough to draw a parallel with microarray experiments performed in a biological research setting, for example when using clinical samples (Gyorffy et al., 2009; Shedden et al., 2005). To address these concerns, we used an additional approach in Paper I, where the performance of various data analysis workflows was evaluated on a cell culture model of hypoxia, where there is sufficient prior knowledge about the anticipated differential expression. The portion of MAQC data used in the current study included three different Affymetrix® microarray platforms (Human Genome U133 Plus 2.0, Human Gene 1.0 ST, Human Exon 1.0 ST) and two independent RNA-seq datasets. A reference for benchmarking was provided by 867 genes, which were evaluated with Taqman® assays.

It is not straightforward to select uniform criteria for benchmarking different methods. Even if a method appears superior based on given criteria and data set, a change in reference data or criteria can lead to different results (Rung and Brazma, 2013). Here, we chose Matthews's correlation coefficient (MCC) as the primary performance indicator, because it is robust to imbalanced classes (i.e. unequal ratio of differentially and equally expressed genes between the samples) and it has been endorsed by MAQC consortium (Consortium et al., 2006). Our analysis of the Taqman® assays indicated such an imbalance when observing 569 differentially expressed and 298 non-differentially expressed genes. In addition, MCC is threshold-dependent, which makes it different from area under the receiver operator curve (AUC), another popular performance metric. We argue that for differential expression, it makes more sense to use MCC at the conventional FDR cutoffs than a threshold-agnostic measure, which ignores the fact that, in practice, differential expression is evaluated at a predetermined FDR level. Although similar in nature, MCC and AUC are distinct from each other in that AUC does not rely on a specific significance level, which is used to reject the null hypothesis (typically 0.05 or 0.01). Instead AUC is calculated from the receiver operating characteristic (ROC) curve over a range of p-values. For example, in cases where false positive rate (FPR) and true positive rate (TPR) are 0 at the specified significance level, yielding an MCC of 0, AUC can still yield favorable values. Using AUC and MCC simultaneously can be useful in situations where the predictions are accurate according to AUC but lack sensitivity at the specified significance level (as indicated by low MCC). This was observed with RankProd, which had favorable AUC values, but low

MCC scores compared to other methods. Ultimately, in order to be comprehensive, we recorded additional performance indicators such as AUC, TPR, FPR, true negative rate (TNR) and false negative rate (FNR).

Preprocessing of MAQC datasets was performed either by RMA (Irizarry et al., 2003b), FARMS (Hochreiter et al., 2006), DFW (Chen et al., 2007) or PLIER (Affymetrix, 2005). Preprocessing, together with differential expression estimation using either Limma (Smyth, 2005) or RankProd (Hong et al., 2006), yielded eight distinct differential expression analysis pipelines that were compared to DEMI using the above mentioned performance metrics. DEMI and RankProd showed little variation in performance metric values across different microarray platforms. DEMI's performance was least favorable on Human Genome U133 Plus 2.0 microarrays where the sensitivity was slightly lower and false negative ratio was higher when compared to other arrays. This might be due to the Human Genome U133 Plus 2.0 microarray being an older design with fewer probes. However, DEMI's performance was nearly indistinguishable between the Human Exon 1.0 ST and Human Gene 1.0 ST arrays.

In addition to exhibiting stable performance across different microarray platforms, DEMI's was also the most accurate when the sample size was very small ($n = 2$). Under these conditions, DEMI demonstrated good performance whereas Limma and RankProd both suffered from a decreased TPR and an increasing FNR i.e. loss of sensitivity. Overall, Limma had the highest sensitivity but also the highest FPR whereas RankProd had almost no false positives that came at a cost of very low sensitivity. Also, it seems that the choice of preprocessing method (RMA, FARMS, DFW, PLIER) had a smaller influence on differential expression analysis results compared to the selection of the differential expression estimation method (Limma, RankProd). PLIER, however, seemed to yield most variable results in comparison to other preprocessing methods.

RNA-seq analysis of MAQC samples did not provide enough advantageous evidence over microarray technology as has been previously claimed (Trapnell et al., 2013). At low mRNA abundance, microarrays are more accurate in detecting differentially expressed genes while RNA-seq produces a systematic false-negative problem that can likely be improved with a higher sequencing depth (Liu et al., 2011; Liu et al., 2014). In addition, although RNA-seq has a wider dynamic range, not restricted by the number of hybridizing molecules as in microarrays, a relatively small number of abundant transcripts can account for the majority of the reads (Bradford et al., 2010). For example Łabaj et al. (2011) have calculated that approximately 75% of measurement power is concentrated on 7% of the known transcriptome. Furthermore, RNA-seq estimates are skewed in favor of longer transcripts that are more likely to produce reads. Hence, there will be more power to detect differential expression in longer transcripts (Bradford et al., 2010). The length bias is consistent with the uniform sampling problem where shorter nucleic acid molecules yield fewer fragments and therefore less reads that can be employed to estimate differential expression. Even when the sequencing depth is increased, genes with low or moderate mRNA levels are difficult to quantify with good precision (Łabaj et al., 2011).

The difference between two RNA populations in MAQC datasets was relatively large with over 38% of the probes on the microarray being differentially expressed. In contrast, in the tissue culture model of hypoxia only 11% of the probes where differentially expressed, although prolonged hypoxia is considered to induce wide-spread changes in gene expression (Lendahl et al., 2009), which was also observed in Paper II. Therefore, we evaluated the nine differential expression analysis pipelines in a cell culture model of hypoxia, which is expected to yield more experimental variability than evident in the RNA pools used by MAQC. Hypoxia response is well characterized and there are several related pathways available in GO database, such as "glycolysis" (GO:0006096) (reviewed in (Lendahl et al., 2009)) and "cellular response to hypoxia" (GO:0071456). Although all methods were able to yield enrichment of these categories among upregulated genes in response to hypoxia, only DEMI and DWF coupled with RankProd performed perfectly with very small sample size ($n = 2$). Similar outcome was observed when upregulated genes in hypoxia were studied for enrichment of HIF-1 and HIF-2 targets in mouse orthologs. This demonstrates the applicability of DEMI even when the number of technical replicates is limited as in pilot studies or clinical settings.

1.2. Application range of DEMI

The possibility to analyze large genomic regions is a novel feature in DEMI, which is difficult to achieve with other differential expression analysis pipelines. We tested it on a dataset consisting of normal prostate epithelial cells (PrEC) and prostate cancer cell line (LNCaP) published by Coolen et al. (2010). More specifically, they studied the suppression of neighboring genes due to chromatin remodeling in cancerous cells and named the process Long-range Epigenetic Silencing (LRES) (Coolen et al., 2010). Although the dataset consisted of only two biological replicates, DEMI successfully identified 80% of downregulated genomic regions as being putative LRES regions proposed in the original paper. This makes DEMI a good choice for studying differential expression on genomic regions to determine common behavior of neighboring genes, which is especially applicable for cancer, where coordinated epigenetic regulation over large regions is characteristic (Coolen et al., 2010). The large number of publicly available cancer microarray datasets provides a good prospect to study such events more extensively.

Last but not least we used DEMI to study gene expression response to varying durations of hypothermia. We identified hypothermia-dependent upregulation of several genes involved in antioxidant response, which might indicate a novel therapeutic route of hypothermia. These results require further validation to pinpoint the exact transcriptional mechanism leading to the observed gene expression response.

DEMI is publicly available as a downloadable R package from CRAN. The package was built so that the users can develop and incorporate custom methods to estimate differential expression on probe-level data. For example, we have

used Wilcoxon-Mann-Whitney rank sum test to estimate differential expression between groups, which is very different to estimating differential expression with Kendall's tau statistic that can evaluate whether the expression levels show a monotonic up- or downward trend. The user is free to incorporate any other method in DEMI analysis pipeline as long as it fits his/her experimental design thus providing a versatile and useful methodology for addressing a large variety of problems (Rung and Brazma, 2013).

1.3. The strengths and limitations of DEMI

A major challenge that microarray analysis workflows face is the summarization of probe-level data on the gene level. Methods such as FARMS (Hochreiter et al., 2006) and DFW (Chen et al., 2007) were specifically designed for such purpose. However due to summarization of probe values into a probe-set value, differential expression estimation will be made on considerably fewer variables, which can have unfavorable effects on statistical power when sample size is small. DEMI resolves this issue by performing differential expression analysis on probe-level before making final inference about the targets e.g. genes. By taking advantage of the high number of simultaneous measurements DEMI preserves more statistical power, enabling it to detect differential expression even when the sample size is very small ($n < 2$). In essence, DEMI is "borrowing" information across targets to make inference about one target (Allison et al., 2006).

DEMI does not make any assumptions on the signal distribution, which in contrary to model-based methods, such as RMA and FARMS, can provide biased estimates when the assumptions are not met (Chen et al., 2007). For example, when estimating background noise, RMA assumes a global background distribution for all probes that is normally distributed (Wu, 2009). In addition, many normalization procedures assume that the biological effect does not alter the shape of the distribution, which might be true if the samples are very similar to each other, but can deviate if the differences in increasing and decreasing target amounts are asymmetrical (Wu, 2009).

Most of the conventional microarray analysis methods rely on predefined Chip Definition File (CDF). As time goes by, these can become outdated, and the probes measuring specific genes become unreliable (Dai et al., 2005). Although DEMI faces similar problems, it currently benefits from the realignment of all probe sequences, independent of the CDF, to the latest versions of the genome, transcriptome and exome. Additionally, when using standard CDF, each probe belongs to a single probe-set whereas in DEMI analysis, a probe can interrogate several distinct targets. Although detrimental to sensitivity, the user is given the choice to specify the number of targets a probe is allowed to interrogate in the DEMI R package. By default, such restrictions are not applied.

In contrast to quantile-normalization, which is a popular method to normalize microarray data, DEMI normalizes probe intensities separately in each sample by relative ranking and does not summarize them on the target level. As

a technical feature, this process is highly parallelizable and therefore scalable to large datasets provided that enough computing power is available. However, a clear benefit of using quantile-normalization together with summarization as used by RMA, FARMS, DFW and PLIER, is that it produces target estimates of expression, which can be used to make arbitrary inferences later. Currently, DEMI does not produce target-specific estimates of expression levels.

Of note, when evaluating differential gene expression in Ngb-deficient mouse retina in response to light-pulse in Paper III, we set an additional cutoff that at least half of the target-specific probes had to be differentially expressed in the same direction. This constraint was enforced, because we wanted to lower the number of false positive findings since the number of differentially expressed probes was relatively small due to the samples being very similar. Even a modest enrichment of on-target differentially expressed probes becomes statistically significant when the background ratio of differentially expressed probes on the microarray is very small. In other words, DEMI's sensitivity grows as the similarity between samples increases. From a practical viewpoint, this restriction made sense, because the primary aim of a microarray experiment is to rank the genes based on observed evidence against the null hypothesis, not just assign a p-value (Smyth, 2004; Smyth et al., 2003). Thus, it makes sense to declare targets as equally expressed, when less than half of the probes indicate differential expression despite the p-value being significant.

From the benchmarking analysis we observed, that RNA-seq was more sensitive in detecting differentially expressed genes. However, this came at a cost of increased false positive findings, resulting in a lower MCC values when compared to the microarray analysis workflows. Although RNA-seq holds huge potential for gene expression studies, compared to microarrays, the reporting standards are almost non-existent (Rung and Brazma, 2013). The distinctive feature of RNA-seq is the uniform sampling of nucleic acid fragments for sequencing. Accordingly, RNA-seq exhibits less accuracy at detecting differential expression of genes with low expression levels, which are less likely to be sampled (Liu et al., 2011; Oshlack and Wakefield, 2009). However the signal intensity of microarrays are proportional only to the expression level of the transcript in addition to the hybridization properties of the probe itself, such as GC content (Dunning et al., 2008; Oshlack and Wakefield, 2009; Wu and Irizarry, 2005). Overall, combining results retrieved by both RNA-seq and microarray technology, will provide a more extensive view on the transcriptional machinery than just pertaining to a single method (Swindell et al., 2014).

Regardless, with better standards for RNA-seq, it has the potential to attain absolute measurements of gene expression levels (Rung and Brazma, 2013), perhaps surpassing microarrays in detection accuracy and providing the opportunity to measure new functional elements in the genome. In principle, RNA-seq retrieves direct counts of mRNA molecules, thus allowing for a wider interpretation of the samples, even between experiments (Wang et al., 2009). Due to practical reasons (i.e. low sampling depth), as for now, this potential is not captured in practice.

1.4. Thoughts on using DEMI for RNA-seq data analysis

In a differential expression analysis, each probe on a microarray describes a genes behavior dependent on the experimental conditions. With a two-group experimental design a probe can be either up- or downregulated or unchanged. Since probes are not an intrinsic feature of an RNA-seq data analysis pipelines the use of conventional microarray data analysis workflows is improbable or even impossible with RNA-seq data. However since RNA-seq data consists of read alignments with specific genomic coordinates, it is possible to count the number of hits a read aligns to a specified genomic location. By dividing the genome or more specifically all the known genes into probes with a pre-determined size, we can count the times a read overlaps with every probe. By utilizing DEMI's methodology we can use the ratios of off-target differentially expressed probes and on-target differentially expressed probes to indicate differential expression of the target gene, similar to microarray analysis. This approach to RNA-seq data analysis requires complex benchmarking. For example, decisions on the best probe length have to be made, or maybe the use of varying probe size depending on gene length is deemed appropriate. A similar argument can be made for extending DEMI to analyze differential expression in proteomics data.

2. Effect of Neuroglobin deficiency

Another important subject addressed in this work, is the effect of Neuroglobin (Ngb) deficiency in mice. Ngb is a neuron-specific globin (Burmester et al., 2000), that was believed to be involved in the oxygen storage and intracellular diffusion of oxygen (Brunori and Vallone, 2007). We conducted a genome wide gene expression study using Affymetrix Mouse Gene ST 1.0 microarrays where we observed an upregulation of several well-established hypoxia responsive genes. Our results confirmed the hypoxia-dependent regulation of pathways that among others included apoptosis, cell growth ("mTOR signaling"), synthesis of ATP ("oxidative phosphorylation") and angiogenesis ("VEGF signaling pathway") in both genotypes. Based on a broader categorization, it can be claimed that global gene expression profiles of Ngb-null and wt mice indicate that chromatin remodeling and mRNA metabolism are among the key regulatory mechanisms that are activated upon prolonged hypoxia.

When quantifying gene expression levels with qPCR, we observed a significant upregulation of *Hif1A* both in Ngb-null and wt mice after 90 minutes of hypoxia, although similar result was only observed for Ngb-null mice when microarrays were used. Hif1A is part of a transcriptional cascade that mediates a broad network of systemic and local responses to hypoxia (Ratcliffe, 2002). In addition, at 24 hours of hypoxia, Ngb-null mice showed a higher expression of *Hif1A* in relation to wt mice, as measured by qPCR.

To further study the potential connection between Ngb and neuroprotection during hypoxic conditions we examined the effect of Ngb deficiency on light-

induced gene expression in the retina (Paper III). Since retina is metabolically one of the most active tissues (Ames, 1992; Anderson and Saltzman, 1964), we anticipated that Ngb deficiency would have a notable effect on light-induced gene expression. More specifically, oxygen consumption rate in retina is dependent on the exposure to light, since dark-adapted retina requires more oxygen compared to light-adapted retina (Ahmed et al., 1993; Linsenmeier, 1986). We speculated that if Ngb was an important oxygen reservoir or if it were important for oxygen metabolism and/or delivery, its deficiency would have a detectable effect on gene expression when oxygen consumption requirements are changed due to exposure to light.

To that end, we conducted a genome wide differential gene expression study using Mouse Exon ST 1.0 arrays, that compared varying durations of light exposure to dark-adapted retina in wt and Ngb-null mice. Our observations indicated, that the effect of light on gene expression was more prevalent than the effect of genotype, which was reflected in the number of differentially expressed genes. To corroborate our results on light exposure, we used DEMI to reanalyze two publicly available datasets (Bedolla and Torre, 2011; Porterfield et al., 2007) from related experiments. We were able to cross-validate light-induced gene expression response between the independent experiments and our study regardless of the genotype. In other words, we saw no indication that Ngb deficiency would lead to an altered light response or to expressional changes indicative of increased amounts of cellular stress caused by decreased oxygen levels. Despite the lack of an obvious link between Ngb and oxygen consumption in retina, we were further interested in the differential expression of marker genes associated with oxygen availability (Semenza, 2012), oxidative stress (Murphy, 2009) and endoplasmatic reticulum (ER)-stress (Xu et al., 2005). Using qPCR, we did not observe any indication of an altered gene expression response between the genotypes, suggesting that Ngb is not important in the normal retinal function upon light exposure.

When ignoring the light effect and pooling experimental treatments together, the differential gene expression analysis between Ngb-null and wt mice revealed several statistically significant genes. The top three differentially expressed genes were *Akap6*, *Entpd4* and *Atp8a2*, of which only *Akap6* was upregulated. Since *Akap6* lies within 35 Mbp of the Ngb locus, its differential expression can be the effect of congenic footprint (Schalkwyk et al., 2007). *Entpd4* and *Atp8a2* were both downregulated, but there is no direct evidence of functional link between these genes and *Ngb*. *Entpd4* is thought to be involved in the rescue of nucleotides from the lysosomal/autophagic vacuole lumen (Biederbick et al., 1999). *Atp8a2* is an adenosine triphosphate (ATP)-dependent lipid flippase that translocates aminophospholipids from the exoplasmic to cytoplasmic leaflets of membranes (Coleman et al., 2014). Additionally, *Atp8a2* is primarily expressed in testes, spinal cord and retina (Cacciagli et al., 2010; Coleman et al., 2009; Zhu et al., 2012) and is necessary for the correct functioning of photoreceptor cells (Coleman et al., 2014), which, curiously, do not express *Ngb* (Hundahl et al., 2012).

CONCLUSIONS

We have designed and implemented a new methodology called Differential Expression from Multiple Indicators (DEMI) to help researchers conduct differential expression analysis using high-density microarrays. Although many well-performing methods have already been popularized, DEMI approaches the data from a new angle, which provides additional benefits not easily achieved with other differential expression analysis workflows. I am convinced that the research community will benefit from a new differential expression analysis workflow for microarrays, even when considering the recent advent of RNA-seq data.

In my thesis I have demonstrated the properties of DEMI and eight other differential gene expression pipelines using a comprehensive benchmarking and a "gold-standard" reference dataset. Although several of the workflows demonstrated good results, DEMI performed equally well with the top performing workflows and was the most stable regardless of the sample size or microarray platform. The ability to correctly identify differentially expressed genes even when there are very few replicates is a valuable property for pilot studies or when samples are hard to obtain, as for example, in clinical cases. Additional benefit comes from DEMI's ability to analyze differential expression on large genomic regions that might encompass several neighboring genes. This feature is especially useful with cancer datasets where gene expression regulation through epigenetic changes encompassing wide regions is characteristic. Although there are many cancer datasets available in the public databases, in most cases they are focused studies looking at individual genes whereas looking neighboring regions together, would provide new insight to the epigenetic modifications caused by chromatin remodeling.

Additionally, DEMI is a self-contained pipeline whereas other methods require to choose from a combination of a preprocessing method and a differential expression estimator. Also, I believe that DEMI's logic of using ratios of differentially expressed on-target and off-target probes is easy to understand and intuitive. More complex analysis can be easily incorporated in DEMI pipeline, as we have demonstrated by using Kendall's tau statistic for time-series analysis. In principle, any differential expression statistic that is concordant with the experimental design could be integrated into DEMI.

We have used DEMI to study whether the notion that Ngb is an important oxygen reservoir in the brain and exhibits neuroprotective capabilities is confirmed. In that case, one would expect the lack of Ngb to have significant consequences under circumstances that require functional oxygen metabolism. By measuring gene expression response to light-pulse in the mouse retina, which is metabolically one of the most active tissues, we did not detect any major differences between Ngb-deficent and wild-type mice. However, we did observe a statistically significant enrichment of differentially expressed genes between our and publicly available datasets that responded similarly to light-pulse. All in all, the present study indicates that Ngb deficiency does not lead to major alterations in light-dependent gene expression response, but leads to subtle systemic differences of currently unknown functional significance.

REFERENCES

Adler, P., Kolde, R., Kull, M., Tkachenko, A., Peterson, H., Reimand, J., and Vilo, J. 2009. Mining for coexpression across hundreds of datasets using novel rank aggregation and visualization methods. *Genome biology* 10:R139.

Affymetrix. 2005. Technical Note. Guide to Probe Logarithmic Intensity Error (PLIER) Estimation.

Ahmed, J., Braun, R.D., Dunn, R., Jr., and Linsenmeier, R.A. 1993. Oxygen distribution in the macaque retina. *Investigative ophthalmology & visual science* 34:516–521.

Alberts, B. 2014. Molecular biology of the cell, Sixth edition. ed.

Allison, D.B., Cui, X., Page, G.P., and Sabripour, M. 2006. Microarray data analysis: from disarray to consolidation and consensus. *Nature reviews. Genetics* 7:55–65.

Ames, A., 3rd. 1992. Energy requirements of CNS cells as related to their function and to their vulnerability to ischemia: a commentary based on studies on retina. *Canadian journal of physiology and pharmacology* 70 Suppl:S158–164.

Anders, S. and Huber, W. 2010. Differential expression analysis for sequence count data. *Genome biology* 11:R106.

Anderson, B., Jr. and Saltzman, H.A. 1964. Retinal Oxygen Utilization Measured by Hyperbaric Blackout. *Archives of ophthalmology* 72:792–795.

Araki, R., Nakahara, M., Fukumura, R., Takahashi, H., Mori, K., Umeda, N., Sujino, M., Inouye, S.T., and Abe, M. 2006. Identification of genes that express in response to light exposure and express rhythmically in a circadian manner in the mouse suprachiasmatic nucleus. *Brain research* 1098:9–18.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25:25–29.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* 129:823–837.

Bedolla, D.E. and Torre, V. 2011. A component of retinal light adaptation mediated by the thyroid hormone cascade. *PloS one* 6:e26334.

Bemmo, A., Benovoy, D., Kwan, T., Gaffney, D.J., Jensen, R.V., and Majewski, J. 2008. Gene expression and isoform variation analysis using Affymetrix Exon Arrays. *BMC genomics* 9:529.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57:289–300.

Benjamini, Y. and Yekutieli, D. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 29:1165–1188.

Bernard, S.A., Gray, T.W., Buist, M.D., Jones, B.M., Silvester, W., Gutteridge, G., and Smith, K. 2002. Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia. *The New England journal of medicine* 346:557–563.

Biederbick, A., Rose, S., and Elsasser, H.P. 1999. A human intracellular apyrase-like protein, LALP70, localizes to lysosomal/autophagic vacuoles. *Journal of cell science* 112 (Pt 15):2473–2484.

Blais, J.D., Filipenko, V., Bi, M., Harding, H.P., Ron, D., Koumenis, C., Wouters, B.G., and Bell, J.C. 2004. Activating transcription factor 4 is translationally regulated by hypoxic stress. *Molecular and cellular biology* 24:7469–7482.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185–193.

Bradford, J.R., Hey, Y., Yates, T., Li, Y., Pepper, S.D., and Miller, C.J. 2010. A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. *BMC genomics* 11:282.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., and Vingron, M. 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature genetics* 29:365–371.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. 2005. Principles of microRNA-target recognition. *PLoS biology* 3:e85.

Brittain, T., Skommer, J., Henty, K., Birch, N., and Raychaudhuri, S. 2010. A role for human neuroglobin in apoptosis. *IUBMB life* 62:878–885.

Brunori, M. and Vallone, B. 2007. Neuroglobin, seven years after. *Cellular and molecular life sciences : CMLS* 64:1259–1268.

Bumgarner, R. 2013. Overview of DNA microarrays: types, applications, and their future. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* Chapter 22:Unit 22 21.

Burmester, T., Weich, B., Reinhardt, S., and Hankeln, T. 2000. A vertebrate globin expressed in the brain. *Nature* 407:520–523.

Cacciagli, P., Haddad, M.R., Mignon-Ravix, C., El-Waly, B., Moncla, A., Missirian, C., Chabrol, B., and Villard, L. 2010. Disruption of the ATP8A2 gene in a patient with a t(10;13) de novo balanced translocation and a severe neurological phenotype. *European journal of human genetics : EJHG* 18:1360–1363.

Chen, R., Khatri, P., Mazur, P.K., Polin, M., Zheng, Y., Vaka, D., Hoang, C.D., Shrager, J., Xu, Y., Vicent, S., Butte, A.J., and Sweet-Cordero, E.A. 2014. A meta-analysis of lung cancer gene expression identifies PTK7 as a survival gene in lung adenocarcinoma. *Cancer research* 74:2892–2902.

Chen, R., Mallelwar, R., Thosar, A., Venkatasubrahmanyam, S., and Butte, A.J. 2008. GeneChaser: identifying all biological and clinical conditions in which genes of interest are differentially expressed. *BMC bioinformatics* 9:548.

Chen, Z., McGee, M., Liu, Q., and Scheuermann, R.H. 2007. A distribution free summarization method for Affymetrix GeneChip arrays. *Bioinformatics* 23:321–327.

Coleman, J.A., Kwok, M.C., and Molday, R.S. 2009. Localization, purification, and functional reconstitution of the P4-ATPase Atp8a2, a phosphatidylserine flippase in photoreceptor disc membranes. *The Journal of biological chemistry* 284:32670–32679.

Coleman, J.A., Zhu, X., Djajadi, H.R., Molday, L.L., Smith, R.S., Libby, R.T., John, S.W., and Molday, R.S. 2014. Phospholipid flippase ATP8A2 is required for normal visual and auditory function and photoreceptor and spiral ganglion cell survival. *Journal of cell science* 127:1138–1149.

Consortium, M. Shi, L. Reid, L.H. Jones, W.D. Shippy, R. Warrington, J.A. Baker, S.C. Collins, P.J. de Longueville, F. Kawasaki, E.S. Lee, K.Y. Luo, Y. Sun, Y.A. Willey, J.C. Setterquist, R.A. Fischer, G.M. Tong, W. Dragan, Y.P. Dix, D.J. Frueh, F.W. Goodsaid, F.M. Herman, D. Jensen, R.V. Johnson, C.D. Lobenhofer, E.K. Puri, R.K. Schrf, U. Thierry-Mieg, J. Wang, C. Wilson, M. Wolber, P.K. Zhang, L. Amur, S. Bao, W. Barbacioru, C.C. Lucas, A.B. Bertholet, V. Boysen, C. Bromley, B. Brown, D. Brunner, A. Canales, R. Cao, X.M. Cebula, T.A. Chen, J.J. Cheng, J. Chu, T.M. Chudin, E. Corson, J. Corton, J.C. Croner, L.J. Davies, C. Davison, T.S. Delenstarr, G. Deng, X. Dorris, D. Eklund, A.C. Fan, X.H. Fang, H. Fulmer-Smentek, S. Fuscoe, J.C. Gallagher, K. Ge, W. Guo, L. Guo, X. Hager, J. Haje, P.K. Han, J. Han, T. Harbottle, H.C. Harris, S.C. Hatchwell, E. Hauser, C.A. Hester, S. Hong, H. Hurban, P. Jackson, S.A. Ji, H. Knight, C.R. Kuo, W.P. LeClerc, J.E. Levy, S. Li, Q.Z. Liu, C. Liu, Y. Lombardi, M.J. Ma, Y. Magnuson, S.R. Maqsodi, B. McDaniel, T. Mei, N. Myklebost, O. Ning, B. Novoradovskaya, N. Orr, M.S. Osborn, T.W. Papallo, A. Patterson, T.A. Perkins, R.G. Peters, E.H. Peterson, R. Philips, K.L. Pine, P.S. Pusztai, L. Qian, F. Ren, H. Rosen, M. Rosenzweig, B.A. Samaha, R.R. Schena, M. Schroth, G.P. Shchegrova, S. Smith, D.D. Staedtler, F. Su, Z. Sun, H. Szallasi, Z. Tezak, Z. Thierry-Mieg, D. Thompson, K.L. Tikhonova, I. Turpaz, Y. Vallanat, B. Van, C. Walker, S.J. Wang, S.J. Wang, Y. Wolfinger, R. Wong, A. Wu, J. Xiao, C. Xie, Q. Xu, J. Yang, W. Zhang, L. Zhong, S. Zong, Y. and Slikker, W., Jr. 2006. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature biotechnology* 24:1151–1161.

Coolen, M.W., Stirzaker, C., Song, J.Z., Statham, A.L., Kassir, Z., Moreno, C.S., Young, A.N., Varma, V., Speed, T.P., Cowley, M., Lacaze, P., Kaplan, W., Robinson, M.D., and Clark, S.J. 2010. Consolidation of the cancer genome into domains of repressive chromatin by long-range epigenetic silencing (LRES) reduces transcriptional plasticity. *Nature cell biology* 12:235–246.

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Martin, F.J., Maurel, T., McLaren, W., Murphy, D.N., Nag, R., Overduin, B., Parker, A., Patricio, M., Perry, E., Pignatelli, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Aken, B.L., Birney, E., Harrow, J., Kinsella, R., Muffato, M., Ruffier, M., Searle, S.M., Spudich, G., Trevanion, S.J., Yates, A., Zerbino, D.R., and Flicek, P. 2014. Ensembl 2015. *Nucleic acids research*.

Dai, M., Wang, P., Boyd, A.D., Kostov, G., Athey, B., Jones, E.G., Bunney, W.E., Myers, R.M., Speed, T.P., Akil, H., Watson, S.J., and Meng, F. 2005. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic acids research* 33:e175.

Denko, N.C., Green, S.L., Edwards, D., and Giaccia, A.J. 2000. p53 checkpoint-defective cells are sensitive to X rays, but not hypoxia. *Experimental cell research* 258:82–91.

DeYoung, M.P., Horak, P., Sofer, A., Sgroi, D., and Ellisen, L.W. 2008. Hypoxia regulates TSC1/2-mTOR signaling and tumor suppression through REDD1-mediated 14-3-3 shuttling. *Genes & development* 22:239–251.

Draghici, S., Khatri, P., Eklund, A.C., and Szallasi, Z. 2006. Reliability and reproducibility issues in DNA microarray measurements. *Trends in genetics : TIG* 22:101–109.

Dudoit, S., Yang, J.Y., Callow, M.J., and Speed, T.P. 2002. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *STATISTICA SINICA* 12:111–139.

Dunning, M.J., Barbosa-Morais, N.L., Lynch, A.G., Tavare, S., and Ritchie, M.E. 2008. Statistical issues in the analysis of Illumina data. *BMC bioinformatics* 9:85.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols* 4:1184–1191.

Ebert, B.L., Firth, J.D., and Ratcliffe, P.J. 1995. Hypoxia and mitochondrial inhibitors regulate expression of glucose transporter-1 via distinct Cis-acting sequences. *The Journal of biological chemistry* 270:29083–29089.

Fago, A., Hundahl, C., Dewilde, S., Gilany, K., Moens, L., and Weber, R.E. 2004. Allosteric regulation and temperature dependence of oxygen binding in human neuroglobin and cytoglobin. Molecular mechanisms and physiological significance. *The Journal of biological chemistry* 279:44417–44426.

Fordel, E., Geuens, E., Dewilde, S., Rottiers, P., Carmeliet, P., Grooten, J., and Moens, L. 2004. Cytoglobin expression is upregulated in all tissues upon hypoxia: an in vitro and in vivo study by quantitative real-time PCR. *Biochemical and biophysical research communications* 319:342–348.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L.J. 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research* 41:D808–815.

Fujita, J. 1999. Cold shock response in mammalian cells. *Journal of molecular microbiology and biotechnology* 1:243–255.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. 2004. affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20:307–315.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y., and Zhang, J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology* 5:R80.

Greer, S.N., Metcalf, J.L., Wang, Y., and Ohh, M. 2012. The updated biology of hypoxia-inducible factor. *The EMBO journal* 31:2448–2460.

Gyorffy, B., Molnar, B., Lage, H., Szallasi, Z., and Eklund, A.C. 2009. Evaluation of microarray preprocessing algorithms based on concordance with RT-PCR in clinical samples. *PloS one* 4:e5645.

Hartemink, A.J., Gifford, D.K., Jaakkola, T.S., and Young, R.A. 2001. Maximum-likelihood estimation of optimal scaling factors for expression array normalization. *Proc. SPIE 4266, Microarrays: Optical Technologies and Informatics, 132 (June 4, 2001)*.

Hedtjarn, M., Mallard, C., Eklind, S., Gustafson-Brywe, K., and Hagberg, H. 2004. Global gene expression in the immature brain after hypoxia-ischemia. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* 24:1317–1332.

Hochreiter, S., Clevert, D.A., and Obermayer, K. 2006. A new summarization method for Affymetrix probe level data. *Bioinformatics* 22:943–949.

Holder, D., Raubertas, R.F., Pikouns, V.B., Svetnik, V., and Soper, K. 2001. Statistical Analysis of high density oligonucleotide arrays: a SAFER approach. *Proceedings of the ASA Annual Meeting*.

Hong, F., Breitling, R., McEntee, C.W., Wittner, B.S., Nemhauser, J.L., and Chory, J. 2006. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* 22:2825–2827.

Hota, K.B., Hota, S.K., Srivastava, R.B., and Singh, S.B. 2012. Neuroglobin regulates hypoxic response of neuronal cells through Hif-1alpha- and Nrf2-mediated mechanism. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism* 32:1046–1060.

Hundahl, C.A., Allen, G.C., Nyengaard, J.R., Dewilde, S., Carter, B.D., Kelsen, J., and Hay-Schmidt, A. 2008a. Neuroglobin in the rat brain: localization. *Neuroendocrinology* 88:173–182.

Hundahl, C.A., Fahrenkrug, J., Luuk, H., Hay-Schmidt, A., and Hannibal, J. 2012. Restricted expression of Neuroglobin in the mouse retina and co-localization with Melanopsin and Tyrosine Hydroxylase. *Biochemical and biophysical research communications* 425:100–106.

Hundahl, C.A., Hannibal, J., Fahrenkrug, J., Dewilde, S., and Hay-Schmidt, A. 2010. Neuroglobin expression in the rat suprachiasmatic nucleus: colocalization, innervation, and response to light. *The Journal of comparative neurology* 518:1556–1569.

Hundahl, C.A., Kelsen, J., Dewilde, S., and Hay-Schmidt, A. 2008b. Neuroglobin in the rat brain (II): co-localisation with neurotransmitters. *Neuroendocrinology* 88:183–198.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. 2003a. Summaries of Affymetrix GeneChip probe level data. *Nucleic acids research* 31:e15.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. 2003b. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264.

Iwakoshi, N.N., Lee, A.H., Vallabhajosyula, P., Otipoby, K.L., Rajewsky, K., and Glimcher, L.H. 2003. Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1. *Nature immunology* 4:321–329.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R., and Shoemaker, D.D. 2003. Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302:2141–2144.

Kanehisa, M. and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28:27–30.

Kothari, S., Cizeau, J., McMillan-Ward, E., Israels, S.J., Bailes, M., Ens, K., Kirshenbaum, L.A., and Gibson, S.B. 2003. BNIP3 plays a role in hypoxic cell death in human epithelial cells that is inhibited by growth factors EGF and IGF. *Oncogene* 22:4734–4744.

Łabaj, P.P., Leparc, G.G., Linggi, B.E., Markillie, L.M., Wiley, H.S., and Kreil, D.P. 2011. Characterization and improvement of RNA-Seq precision in quantitative transcript expression profiling. *Bioinformatics* 27:i383–391.

Lendahl, U., Lee, K.L., Yang, H., and Poellinger, L. 2009. Generating specificity and diversity in the transcriptional response to hypoxia. *Nature reviews. Genetics* 10:821–832.

Linker, R.A., Lee, D.H., Ryan, S., van Dam, A.M., Conrad, R., Bista, P., Zeng, W., Hronowsky, X., Buko, A., Chollate, S., Ellrichmann, G., Bruck, W., Dawson, K., Goelz, S., Wiese, S., Scannevin, R.H., Lukashev, M., and Gold, R. 2011. Fumaric acid esters exert neuroprotective effects in neuroinflammation via activation of the Nrf2 antioxidant pathway. *Brain : a journal of neurology* 134:678–692.

Linsenmeier, R.A. 1986. Effects of light and darkness on oxygen distribution and consumption in the cat retina. *The Journal of general physiology* 88:521–542.

Liu, S., Lin, L., Jiang, P., Wang, D., and Xing, Y. 2011. A comparison of RNA-Seq and high-density exon array for detecting differential gene expression between closely related species. *Nucleic acids research* 39:578–588.

Liu, Y., Zhou, J., and White, K.P. 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics* 30:301–304.

Lu, X. and Kang, Y. 2010. Hypoxia and hypoxia-inducible factors: master regulators of metastasis. *Clinical cancer research: an official journal of the American Association for Cancer Research* 16:5928–5935.

Mendez, O., Zavadil, J., Esencay, M., Lukyanov, Y., Santovasi, D., Wang, S.C., Newcomb, E.W., and Zagzag, D. 2010. Knock down of HIF-1alpha in glioma cells reduces migration in vitro and invasion in vivo and impairs their ability to form tumor spheres. *Molecular cancer* 9:133.

Metzker, M.L. 2010. Sequencing technologies – the next generation. *Nature reviews. Genetics* 11:31–46.

Miller, M.B. and Tang, Y.W. 2009. Basic concepts of microarrays and potential applications in clinical microbiology. *Clinical microbiology reviews* 22:611–633.

Miyazawa, T., Tamura, A., Fukui, S., and Hossmann, K.A. 2003. Effect of mild hypothermia on focal cerebral ischemia. Review of experimental studies. *Neurological research* 25:457–464.

Murphy, B.J., Laderoute, K.R., Chin, R.J., and Sutherland, R.M. 1994. Metallothionein IIA is up-regulated by hypoxia in human A431 squamous carcinoma cells. *Cancer research* 54:5808–5810.

Murphy, M.P. 2009. How mitochondria produce reactive oxygen species. *The Biochemical journal* 417:1–13.

Nacu, S., Yuan, W., Kan, Z., Bhatt, D., Rivers, C.S., Stinson, J., Peters, B.A., Modrusan, Z., Jung, K., Seshagiri, S., and Wu, T.D. 2011. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. *BMC medical genomics* 4:11.

Nishiyama, H., Itoh, K., Kaneko, Y., Kishishita, M., Yoshida, O., and Fujita, J. 1997. A glycine-rich RNA-binding protein mediating cold-inducible suppression of mammalian cell growth. *The Journal of cell biology* 137:899–908.

Oshlack, A. and Wakefield, M.J. 2009. Transcript length bias in RNA-seq data confounds systems biology. *Biology direct* 4:14.

Polderman, K.H. 2008. Induced hypothermia and fever control for prevention and treatment of neurological injuries. *Lancet* 371:1955–1969.

Porterfield, V.M., Piontkivska, H., and Mintz, E.M. 2007. Identification of novel light-induced genes in the suprachiasmatic nucleus. *BMC neuroscience* 8:98.

Pradervand, S., Paillusson, A., Thomas, J., Weber, J., Wirapati, P., Hagenbuchle, O., and Harshman, K. 2008. Affymetrix Whole-Transcript Human Gene 1.0 ST array is highly concordant with standard 3' expression arrays. *BioTechniques* 44:759–762.

Qin, L.X., Beyer, R.P., Hudson, F.N., Linford, N.J., Morris, D.E., and Kerr, K.F. 2006. Evaluation of methods for oligonucleotide array data via quantitative real-time PCR. *BMC bioinformatics* 7:23.

Ratcliffe, P.J. 2002. From erythropoietin to oxygen: hypoxia-inducible factor hydroxylases and the hypoxia signal pathway. *Blood purification* 20:445–450.

Reimand, J., Arak, T., and Vilo, J. 2011. g:Profiler – a web server for functional interpretation of gene lists (2011 update). *Nucleic acids research* 39:W307–315.

Relogio, A., Schwager, C., Richter, A., Ansorge, W., and Valcarcel, J. 2002. Optimization of oligonucleotide-based DNA microarrays. *Nucleic acids research* 30:e51.

Rice, J.C. and Allis, C.D. 2001. Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Current opinion in cell biology* 13:263–273.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

Rodriguez-Enriquez, S., Carreno-Fuentes, L., Gallardo-Perez, J.C., Saavedra, E., Quezada, H., Vega, A., Marin-Hernandez, A., Olin-Sandoval, V., Torres-Marquez, M.E., and Moreno-Sanchez, R. 2010. Oxidative phosphorylation is impaired by prolonged hypoxia in breast and possibly in cervix carcinoma. *The international journal of biochemistry & cell biology* 42:1744–1751.

Rung, J. and Brazma, A. 2013. Reuse of public genome-wide gene expression data. *Nature reviews. Genetics* 14:89–99.

Rustici, G., Kolesnikov, N., Brandizi, M., Burdett, T., Dylag, M., Emam, I., Farne, A., Hastings, E., Ison, J., Keays, M., Kurbatova, N., Malone, J., Mani, R., Mupo, A., Pedro Pereira, R., Pilicheva, E., Rung, J., Sharma, A., Tang, Y.A., Ternent, T., Tikhonov, A., Welter, D., Williams, E., Brazma, A., Parkinson, H., and Sarkans, U. 2013. ArrayExpress update – trends in database growth and links to data analysis tools. *Nucleic acids research* 41:D987–990.

Sartor, M., Medvedovic, M., and Aronow, B. 2003. Microarray Data Normalization: The Art and Science of Overcoming Technical Variance to Maximize the Detection of Biologic Variance. *In* A Beginner's Guide to Microarrays (E. Blalock, ed.) pp. 151–178. Springer US.

Schalkwyk, L.C., Fernandes, C., Nash, M.W., Kurrikoff, K., Vasar, E., and Koks, S. 2007. Interpretation of knockout experiments: the congenic footprint. *Genes, brain, and behavior* 6:299–303.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467–470.

Schmidt, M., Giessl, A., Laufs, T., Hankeln, T., Wolfrum, U., and Burmester, T. 2003. How does the eye breathe? Evidence for neuroglobin-mediated oxygen supply in the mammalian retina. *The Journal of biological chemistry* 278:1932–1935.

Schödel, J., Oikonomopoulos, S., Ragoussis, J., Pugh, C.W., Ratcliffe, P.J., and Mole, D.R. 2011. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood* 117:e207–217.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. 2011. Global quantification of mammalian gene expression control. *Nature* 473:337–342.

Semenza, G.L. 2012. Hypoxia-inducible factors in physiology and medicine. *Cell* 148:399–408.

Shedden, K., Chen, W., Kuick, R., Ghosh, D., Macdonald, J., Cho, K.R., Giordano, T.J., Gruber, S.B., Fearon, E.R., Taylor, J.M., and Hanash, S. 2005. Comparison of seven methods for producing Affymetrix expression scores based on False Discovery Rates in disease profiling data. *BMC bioinformatics* 6:26.

Slonim, D.K. 2002. From patterns to pathways: gene expression data analysis comes of age. *Nature genetics* 32 Suppl:502–508.

Smyth, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* 3:Article3.

Smyth, G.K. 2005. Limma: linear models for microarray data. *Bioinformatics and Computational Biology Solutions Using {R} and BioConductor* 397–420.

Smyth, G.K., Yang, Y.H., and Speed, T. 2003. Statistical issues in cDNA microarray data analysis. *Methods in molecular biology* 224:111–136.

Stratowa, C. xps: Processing and Analysis of Affymetrix Oligonucleotide Arrays including Exon Arrays, Whole Genome Arrays and Plate Arrays. *R package version 1.22.2*.

Sun, Y., Jin, K., Mao, X.O., Zhu, Y., and Greenberg, D.A. 2001. Neuroglobin is up-regulated by and protects neurons from hypoxic-ischemic injury. *Proceedings of the National Academy of Sciences of the United States of America* 98:15306–15311.

Swindell, W.R., Xing, X., Voorhees, J.J., Elder, J.T., Johnston, A., and Gudjonsson, J.E. 2014. Integrative RNA-seq and microarray data analysis reveals GC content and gene length biases in the psoriasis transcriptome. *Physiological genomics* 46:533–546.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology* 31:46–53.

Trapnell, C., Pachter, L., and Salzberg, S.L. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111.

van't Veer, L.J. and Bernards, R. 2008. Enabling personalized cancer medicine through analysis of gene-expression patterns. *Nature* 452:564–570.

Wang, A., Tang, Z., Park, I.H., Zhu, Y., Patel, S., Daley, G.Q., and Li, S. 2011. Induced pluripotent stem cells for neural tissue engineering. *Biomaterials* 32:5023–5032.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476.

Wang, Z., Gerstein, M., and Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* 10:57–63.

Wellmann, S., Bettkober, M., Zelmer, A., Seeger, K., Faigle, M., Eltzschig, H.K., and Buhrer, C. 2008. Hypoxia upregulates the histone demethylase JMJD1A via HIF-1. *Biochemical and biophysical research communications* 372:892–897.

Wu, Z. 2009. A review of statistical methods for preprocessing oligonucleotide microarrays. *Statistical methods in medical research* 18:533–541.

Wu, Z. and Irizarry, R.A. 2005. Stochastic models inspired by hybridization theory for short oligonucleotide arrays. *Journal of computational biology : a journal of computational molecular cell biology* 12:882–893.

Xu, C., Bailly-Maitre, B., and Reed, J.C. 2005. Endoplasmic reticulum stress: cell life and death decisions. *The Journal of clinical investigation* 115:2656–2664.

Yun, Z., Maecker, H.L., Johnson, R.S., and Giaccia, A.J. 2002. Inhibition of PPAR gamma 2 gene expression by the HIF-1-regulated gene DEC1/Stra13: a mechanism for regulation of adipogenesis by hypoxia. *Developmental cell* 2:331–341.

Zhao, S., Fung-Leung, W.P., Bittner, A., Ngo, K., and Liu, X. 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PloS one* 9:e78644.

Zhijin, W., Irizarry, R.A., Gentleman, R., Martinez-Murillo, F., and Spencer, F. 2004. A model-based background adjustment for oligonucleotide expression arrays. *Journal of the American statistical Association* 99:909–917.

Zhu, X., Libby, R.T., de Vries, W.N., Smith, R.S., Wright, D.L., Bronson, R.T., Seburn, K.L., and John, S.W. 2012. Mutations in a P-type ATPase gene cause axonal degeneration. *PLoS genetics* 8:e1002853.

APPENDIX I

An example of an output using differential expression analysis with DEMI.

clusterID	targetID	POTIC	POT	PIC	PT	P.value	FDR	geneSymbol	Description
BRAIN[H]_UHR	ENSG00000008710	982	1388	349880	1952593	0	0	*PKD1*	polycystic kidney disease 1 (autoso ...
BRAIN[H]_UHR	ENSG00000131018	716	895	349880	1952593	0	0	*SYNE1*	spectrin repeat containing, nuclear ...
BRAIN[H]_UHR	ENSG00000178104	699	808	349880	1952593	0	0	*PDE4DIP*	phosphodiesterase 4D interacting ...
BRAIN[H]_UHR	ENSG00000224078	677	771	349880	1952593	0	0	*SNHG14*	small nucleolar RNA host gene 14 ...
BRAIN[H]_UHR	ENSG00000264708	512	544	349880	1952593	0	0	*PDE4DIP*	Phosphodiesterase 4D interacting ...
BRAIN[H]_UHR	ENSG00000170160	510	583	349880	1952593	1.04E-293	1.29E-289	*CCDC144A*	coiled-coil domain containing 144A
BRAIN[L]_UHR	ENSG00000013573	476	541	389277	1952593	1.89E-255	2.02E-251	*DDX11*	DEAD/H (Asp-Glu-Ala-Asp/His) box ...
BRAIN[L]_UHR	ENSG00000143341	456	506	389277	1952593	2.36E-255	2.20E-251	*HMCN1*	hemicentin 1
BRAIN[H]_UHR	ENSG00000198626	412	464	349880	1952593	1.98E-243	1.64E-239	*RYR2*	ryanodine receptor 2 (cardiac)
BRAIN[H]_UHR	ENSG00000168702	398	447	349880	1952593	2.78E-236	2.08E-232	*LRP1B*	low density lipoprotein ...
BRAIN[H]_UHR	ENSG00000183889	840	1543	349880	1952593	3.11E-228	2.11E-224	*PKD1P1*	NPIP-like protein 1
BRAIN[L]_UHR	ENSG00000153107	513	656	389277	1952593	6.33E-226	3.94E-222	*ANAPC1*	anaphase promoting complex ... ...
BRAIN[H]_UHR	ENSG00000173064	397	466	349880	1952593	1.48E-219	8.53E-216	*HECTD4*	HECT domain containing E3 ...
BRAIN[H]_UHR	ENSG00000144406	330	351	349880	1952593	1.66E-215	8.87E-212	*UNC80*	unc-80 homolog (C. elegans)
BRAIN[H]_UHR	ENSG00000189056	308	319	349880	1952593	7.87E-212	3.92E-208	*RELN*	reelin
...	...	...	...	...	...	...	...	...	...

POTIC, probes on target in the cluster, represents total number of on-target probes that are differentially expressed; POT, probes on target, represents total number of on-target probes, including differentially and non-differentially expressed probes; PIC, probes in cluster, represents total number of differentially expressed probes; PT, probes total, represents total number of probes interrogated by the microarray.

SUMMARY IN ESTONIAN

Diferentsiaalse geeniekspressiooni erinevuste hindamine mitmete indikaatorite alusel

Geenide aktiivsus erinevates rakutüüpides varieerub. Geeniaktiivsuse hoolikas ja kontrollitud reguleerimine võimaldab rakkudel saavutada, säilitada ja muuta oma funktsiooni ja füsioloogiat, aidates ühtlasi kohaneda muutuva keskkonnaga. Geeniaktiivsus ehk geeniekspressioon on eelkõige määratletud kui mRNA hulk rakus, mida tänapäeval on võimalik mitme erineva tehnoloogia abil mõõta. Alates 1990-ndate keskpaigast on geeniekspressiooni mõõtmisel teiste meetodite kõrval kasutatud ka mikrokiibi tehnoloogiat, mis võimaldab mõõta ekspressiooni paljudel geenidel üheaegselt. Tänaseks on mikrokiibi andmete hulk märkimisväärselt kasvanud ning välja on töötatud mitmeid lahendusi diferentsiaalse ekspressiooni analüüsimiseks ning standardeid andmete salvestamiseks avalikes andmebaasides. Enamasti on neid andmeid vaadeldud ühe uuringu piiratud vaatenurgast, mistõttu avalike andmete kasutamine ja analüüside kordamine alternatiivsetele küsimustele vastamiseks leiab aina suuremat kasutust.

Antud töö on keskendunud uue diferentsiaalse ekspressiooni analüüsi meetodi väljaarendamisele ja selle võrdlemisele olemasolevate meetoditega. Meetod *Differential Expression from Multiple Indicators* (DEMI) kasutab ära sondide arvukust kõrg-tihedusega mikrokiipidel ning on võimeline hindama diferentsiaalset geeniekspressiooni ka väheste replikaatide olemasolu korral. Lisaks võimaldab DEMI mikrokiipide abil analüüsida suuri genoomi piirkondi, pakkudes sedasi ülevaadet epigeneetilistest kromatiini modifikatsioonidest tingitud laialdastest ekspressiooni muutustest, mis võib enda alla haarata ka lähedalasuvaid geene. Viimast analüüsi on võimalik teha ka traditsioonilisi mikrokiibi analüüsi meetodeid kasutades, kuid see on üldjuhul keerulisem ja aeganõudvam ning vajab lisaks diferentsiaalse ekspressiooni analüüsile ühte täiendavat etappi.

Võrdlesime DEMI-t kaheksa alternatiivse mikrokiibi diferentsiaalse ekspressiooni analüüsi meetodiga, kasutades *MicroArray Quality Control* (MAQC) konsortsiumi proove. Võrdluses kasutasime veel RNA-seq-i andmestikke ja analüüsi meetodeid ning positiivse kontrollina MAQC proovidel tehtud umbes 850 geeni kvantitatiivse polümeraasi ahelreaktsiooni (qPCR) mõõtmisi. Meetodi headust hinnati Matthewsi korrelatsioonikordaja (MCC), kõvera-aluse pindala (AUC) ning tundlikkuse ja spetsiifilisuse näitajate abil. Antud näitajate alusel oli DEMI sooritus pidevalt parimate hulgas ning kõige stabiilsem, sealhulgas olukordades, kus tehnilisi või bioloogilisi replikaate oli väga vähe ($n = 2$). Lisaks oli DEMI sooritus stabiilne sõltumata mikrokiibi platvormist ja varieerus vähem kui alternatiivsed mikrokiibi diferentsiaalse ekspressiooni meetodid, mis sõltusid eelkõige diferentsiaalse ekspressiooni hindamise, ja vähem normaliseerimise, meetodi valikust. RNA-seq-i tundlikkus oli suurem

kui mikrokiipidel, kuid sellega kaasnes ka suurem valepositiivsete arv, mistõttu MCC väärtused olid mikrokiipidega võrreldes madalamad.

DEMI raamistik on rakendatud R-i paketina, mis on internetist vabalt allalaetav (http://cran.r-project.org/web/packages/demi/index.html). Käesolevas doktoritöös kasutati DEMI-t kolmes artiklis geenide diferentsiaalse ekspressiooni hindamiseks. Kasutasime DEMI R-i paketti kaugele-ulatuva epigeentilise vaigistamise (ingl. k. *long-range epigenetic silencing*, LRES) uurimiseks vähi andmestikul, mis publitseeriti 2010 aastal Coolen et al. poolt. Kasutades ainult kahte replikaati grupi kohta, suutis DEMI edukalt tuvastada 38 genoomi piirkonda originaalis avaldatud 47-st, mis olid kromatiini remodelleerimise tõttu vähikoes alla reguleeritud. Lisaks eelnevale rakendasime DEMI-t, et analüüsida hüpotermiast tingitud geeniekspressiooni muutusi erinevates ajapunktides hiire embrüonaalsetes fibroblastides. Uudse tulemusena leidsime, et hüpotermia tingimustes on mitmete antioksüdatiivsete geenide ekspressioon suurenenud. Kasutasime DEMI-t ka Neuroglobiini (Ngb) geeni funktsiooni uurimiseks Ngb-puudulikel hiirtel. Varem on arvatud, et Ngb on oluline hapniku hoidla ajus ning omab olulist rolli sealses hapniku metabolismis. Meie tulemused näitavad, et Ngb puudulikkusest tingitud genotüübi efekt on üsna väike, kui arvestada geeniekspressiooni erinevusi, mis on tingitud madalast hapniku tasemest keskkonnas või vähenenud hapniku tarbimisest hiire reetinas, mis on eksponeeritud valgusele.

Antud doktoritöö demonstreerib DEMI võimalusi erinevate eksperimentaalsete probleemide lahendamisel, kus on võimalik analüüsida diferentsiaalset geeniekspressiooni kõrg-tihedusega mikrokiipidel. Kuigi varasemalt on mitmeid diferentsiaalse geeniekspressiooni analüüsi meetodeid juba välja töötatud ning aina suuremat populaarsust on kogunud diferentsiaalse ekspressiooni hindamine RNA-seq-i abil, siis vaatamata sellele, ma usun, et DEMI pakub teadlaskonnale piisavalt lisandväärtust ja paindlikkust, et seda oma andmete analüüsimisel kasutada.

ACKNOWLEDGEMENT

I would like to thank my supervisors Eero Vasar and Jaak Vilo for their support and guidance throughout all these years. Futhermore, I am eternally grateful to my supervisor Hendrik Luuk, who not only is an unlimited source of exemplary rational and invigorating insight but also an inspiration and a very good friend. Thank you Hendrik!

I am also thankful for the peer-reviewers Viljar Jaks and Martti Laan for their valuable feedback and useful suggestions. I bow before my work colleagues from Ümarlaud and BIIT and my course mates for making the working hours and also after hours fun and for creating an environment where one can shamelessly express ones ideas. In that regard, my special thank goes to Christian Ansgar Hundahl, Kristjan Korjus, Raivo Kolde, Jüri Reimand, Hedi Peterson, Priit Adler, Tanel Visnapuu, Riin Reimets, Kattri-Liis Eskla, Tanel Traks, Paula Reemann, Karin Tein, Veiko Voolaid and Margus Putku.

Last but not least I would like to thank my mom and my dad, my brother Martin and Ulla for always supporting me in any situation. You are the cornerstone of my life.

ORIGINAL PUBLICATIONS

CURRICULLUM VITAE

Name: Sten Ilmjärv
Date of birth: 02.06.1985, Tallinn, Estonia
Citizenship: Estonian
E-mail: sten.ilmjarv@gmail.com

Formal education:

2009–	Neuroscience PhD program, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia
2007–2009	MSc in gene technology, University of Tartu, Estonia. The title of my thesis was "Researching promoter sequences on the basis of gene expression dynamics in mouse BxD lines".
2003–2007	BSc in gene technology, University of Tartu, Estonia. The title of my thesis was "Describing transcription starts sites for *Saccharomyces Cerevisiae*".
2006–2007	ERASMUS semester in University of Bielefeld, Germany
2000–2003	Mathematics program for Gymnasium, The Gifted and Talented Development Centre, University of Tartu, Estonia
1992–2003	Õismäe Humanitaargymnasium, Tallinn, Estonia
1991–1992	Anija elementary school, Estonia

Professional experience:

2014–	Research Assistant, Group of Prof. Karl-Heinz Krause, Department of Pathology and Immunology, Centre Medical Universitaire, Geneva, Switzerland
2013–2014	Bioinformatic specialist, Institute of Biomedicine and Translational Medicine, Faculty of Medicine, University of Tartu, Estonia
2009–2014	Bioinformatician (0.5), Quretec Ltd.
2004–2014	Bioinformatics and info technology workgroup (BIIT, http://biit.cs.ut.ee/) at University of Tartu, Estonia
2003–2004	Member of the finance team in a value conference "To be or to have" organized by Estonian National Youth Council

Teaching experience:

2012/2013	Lecturer in *Bioinformatics* course for computer scientists at
2011/2012	University of Tartu, Estonia

List of publications

1. Reemann, Paula; Reimann, Ene; **Ilmjärv**, **Sten**; Porosaar, Orm; Silm, Helgi; Jaks, Viljar; Kingo, Külli; Kõks, Sulev (2014). *Melanocytes In the Skin – Comparative Whole Transcriptome Analysis of Main Skin Cell Types*. In Press in Plos ONE.
2. **Ilmjärv**, **Sten**; Reimets, Riin; Hundahl, Christian Ansgar; Luuk, Hendrik (2014). *Effect of light on global gene expression in the neuroglobin-deficient mouse retina*. Biomedical Reports, 2(6): 780–786.
3. **Ilmjärv**, **Sten**; Hundahl, Christian Ansgar; Reimets, Riin; Niitsoo, Margus; Kolde, Raivo; Vilo, Jaak; Vasar, Eero; Luuk, Hendrik (2014). *Estimating differential expression from multiple indicators*. Nucleic Acid Research, 42 (8): e72.
4. Hundahl, Christian Ansgar; Luuk, Hendrik; **Ilmjärv**, **Sten**; Falktoft, Birgitte; Raida, Zindy; Vikesaa, Jonas; Friis-Hansen, Lennart; Hay-Schmidt, Anders (2011). *Neuroglobin-Deficiency Exacerbates Hif1A and c-FOS Response, but Does Not Affect Neuronal Survival during Severe Hypoxia In Vivo*. PLoS ONE, 6(12, e28160), 1–14.
5. Kuegler, Philipp B.; Zimmer, Bastian; Waldmann, Tanja; Baudis, Birte; **Ilmjärv**, **Sten**; Hescheler, Jürgen; Gaughwin, Phil; Brundin, Patrik; Mundy, William; Bal-Price, Anna K.; Schrattenholz, André; Krause, Karl-Heinz; van Thriel, Christoph; Rao, Mahendra S.; Kadereit, Suzanne; Leist, Marcel (2010). *Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing*. ALTEX, 27(1), 17–42.

ELULOOKIRJELDUS

Nimi: Sten Ilmjärv
Sünniaeg: 02.06.1985, Tallinn, Eesti
Kodakondsus: Eesti
E-mail: sten.ilmjarv@gmail.com

Haridus:

2009 –	Neuroteaduste PhD programm, Bio- ja Siirdemeditsiini Instituut, Arstiteaduskond, Tartu Ülikool
2007 – 2009	MSc geenitehnoloogias, Tartu Ülikool. Magistritöö pealkiri oli "Promootorjärjestuste uurimine BxD hiireliini geeniekspressiooni dünaamika alusel".
2003 – 2007	BSc geenitehnoloogias, Tartu Ülikool. Bakalaureuse töö pealkiri oli "Transkriptsiooni algussaitide kirjeldamine pagaripärmil".
2006 – 2007	Erasmuse vahetusüliõpilane, Bielefeldi Ülikool, Saksamaa
2000 – 2003	Tartu Ülikooli Täppisteaduste Kool, matemaatika eriala
1992 – 2003	Õismäe Humanitaargümnaasium, Tallinn
1991 – 1992	Anija algkool

Töökogemus:

2014 –	Bioinformaatika assistent, Prof. Karl-Heinz Krause grupp, Patoloogia ja Immunoloogia Instituut, Meditsiini teaduskond, Genf, Šveits
2013 – 2014	Bioinformaatika spetsialist, Bio- ja Siirdemeditsiini Instituut, Arstiteaduskond, Tartu Ülikool
2009 – 2014	Bioinformaatik (0.5), Quretec
2004 – 2014	Bioinformaatika ja infotehnoloogia töögrupp (BIIT, http://biit.cs.ut.ee/), Tartu Ülikool
2003 – 2004	Väärtuskonverentsi „Olla või omada", mida korraldas Eesti Noorteühenduste Liit, finantstoimkonna liige

Õpetamise kogemus:

2012/2013	*Bioinformaatika* kursuse õppejõud arvutiteadlastele, Tartu
2011/2012	Ülikool

Nimekiri avaldatud publikatsioonidest:

1. Reemann, Paula; Reimann, Ene; **Ilmjärv**, **Sten**; Porosaar, Orm; Silm, Helgi; Jaks, Viljar; Kingo, Külli; Kõks, Sulev (2014). *Melanocytes In the Skin – Comparative Whole Transcriptome Analysis of Main Skin Cell Types*. In Press in Plos ONE.
2. **Ilmjärv**, **Sten**; Reimets, Riin; Hundahl, Christian Ansgar; Luuk, Hendrik (2014). *Effect of light on global gene expression in the neuroglobin-deficient mouse retina*. Biomedical Reports, 2(6): 780–786.
3. **Ilmjärv**, **Sten**; Hundahl, Christian Ansgar; Reimets, Riin; Niitsoo, Margus; Kolde, Raivo; Vilo, Jaak; Vasar, Eero; Luuk, Hendrik (2014). *Estimating differential expression from multiple indicators*. Nucleic Acid Research, 42 (8): e72.
4. Hundahl, Christian Ansgar; Luuk, Hendrik; **Ilmjärv**, **Sten**; Falktoft, Birgitte; Raida, Zindy; Vikesaa, Jonas; Friis-Hansen, Lennart; Hay-Schmidt, Anders (2011). *Neuroglobin-Deficiency Exacerbates Hif1A and c-FOS Response, but Does Not Affect Neuronal Survival during Severe Hypoxia In Vivo*. PLoS ONE, 6(12, e28160), 1–14.
5. Kuegler, Philipp B.; Zimmer, Bastian; Waldmann, Tanja; Baudis, Birte; **Ilmjärv**, **Sten**; Hescheler, Jürgen; Gaughwin, Phil; Brundin, Patrik; Mundy, William; Bal-Price, Anna K.; Schrattenholz, André; Krause, Karl-Heinz; van Thriel, Christoph; Rao, Mahendra S.; Kadereit, Suzanne; Leist, Marcel (2010). *Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicity testing*. ALTEX, 27(1), 17–42.

DISSERTATIONES NEUROSCIENTIAE UNIVERSITATIS TARTUENSIS

1. **Sirli Raud.** Cholecystokinin$_2$ receptor deficient mice: changes in function of GABA-ergic system. Tartu, 2005.
2. **Kati Koido.** Single-nucleotide polymorphism profiling of 22 candidate genes in mood and anxiety disorders. Tartu, 2005.
3. **Dzhamilja Safiulina.** The studies of mitochondria in cultured cerebellar granule neurons: characterization of mitochondrial function, volume homeostasis and interaction with neurosteroids. Tartu, 2006.
4. **Tarmo Areda.** Behavioural and neurogenetic study of mechanisms related to cat odour induced anxiety in rodents. Tartu, 2006.
5. **Aleksei Nelovkov.** Behavioural and neurogenetic study of molecular mechanisms involved in regulation of exploratory behaviour in rodents. Tartu, 2006.
6. **Annika Vaarmann**. The studies on cystatin B deficient mice: neurochemical and behavioural alterations in animal model of progressive myoclonus epilepsy of Unverricht-Lundborg type. Tartu, 2007.
7. **Urho Abramov.** Sex and environmental factors determine the behavioural phenotype of mice lacking CCK_2 receptors: implications for the behavioural studies in transgenic lines. Tartu, 2008.
8. **Hendrik Luuk.** Distribution and behavioral effects of WFS1 protein in the central nervous system. Tartu, 2009.
9. **Anne Must.** Studies on molecular genetics of male completed suicide in Estonian population. Tartu, 2009.
10. **Kaido Kurrikoff**. Involvement of cholecystokinin in chronic pain mechanisms and endogenous antinociception. Tartu, 2009.
11. **Anu Aonurm-Helm.** Depression-like phenotype and altered intracellular signalling in neural cell adhesion molecule (NCAM)-deficient mice. Tartu, 2010.
12. **Silva Sütt.** Role of endocannabinoid system and *Wfs1* in regulation of emotional behaviour: behavioural, pharmacological and genetic studies. Tartu, 2010.
13. **Mari-Anne Philips.** Characterization of *Myg1* gene and protein: expression patterns, subcellular localization, gene deficient mouse and functional polymorphisms in human. Tartu, 2010.
14. **Ranno Rätsep.** Genetics of psoriasis and vitiligo, focus on IL10 family cytokines. Tartu, 2010.
15. **Kairit Joost.** Selective screening of metabolic diseases in Estonia: the application of new diagnostic methods. Tartu, 2012, 143 p.
16. **Monika Jürgenson.** A complex phenotype in mice with partial or complete deficiency of the NCAM protein. Tartu, 2012, 117 p.

17. **Ene Reimann.** Description of the cytokines and cutaneous neuroendocrine system in the development of vitiligo. Tartu, 2012, 117 p.
18. **Jürgen Innos.** Behavioural, pharmacological and neurochemical characterisation of limbic system-associated membrane protein (LSAMP) deficient mice. Tartu, 2013, 113 p.
19. **Kaili Anier.** The role of DNA methylation in the development of cocaine-induced behavioural sensitisation. Tartu, 2013, 147 p.
20. **Maarika Liik.** Cognitive functioning, perceived cognition, subjective complaints and symptoms of depression in patients with epilepsy: neuropsychological assessment and spet brain imaging study. Tartu, 2014, 124 p.