



## RNA sequencing of archived neonatal dried blood spots



Jonas Bybjerg-Grauholm<sup>a,b,\*</sup>, Christian Munch Hagen<sup>a,b,1</sup>, Sok Kean Khoo<sup>c</sup>, Maria Louise Johannesen<sup>a,b</sup>, Christine Søholm Hansen<sup>a,b</sup>, Marie Bækvad-Hansen<sup>a,b</sup>, Michael Christiansen<sup>a,b,d</sup>, David Michael Hougaard<sup>a,b</sup>, Mads V. Hollegaard<sup>a,b</sup>

<sup>a</sup> Department of Congenital Disorders, Statens Serum Institut, Copenhagen DK-2300, Denmark

<sup>b</sup> iPSYCH, The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Denmark

<sup>c</sup> Department of Cell and Molecular Biology, Grand Valley State University, Grand Rapids, MI 49503, USA

<sup>d</sup> Department of Biomedicine, University of Copenhagen, Copenhagen N DK-2200, Denmark

### ARTICLE INFO

#### Article history:

Received 25 October 2016

Received in revised form 16 December 2016

Accepted 16 December 2016

Available online 24 December 2016

#### Keywords:

Gene expression

RNA-seq

Limited material

Neonatal Screening

Guthrie Cards

Dried blood spots

### ABSTRACT

Neonatal dried blood spots (DBS) are routinely collected on standard Guthrie cards for all-comprising national newborn screening programs for inborn errors of metabolism, hypothyroidism and other diseases. In Denmark, the Guthrie cards are stored at  $-20^{\circ}\text{C}$  in the Danish Neonatal Screening Biobank and each sample is linked to elaborate social and medical registries. This provides a unique biospecimen repository to enable large population research at a perinatal level. Here, we demonstrate the feasibility to obtain gene expression data from DBS using next-generation RNA sequencing (RNA-seq). RNA-seq was performed on five males and five females. Sequencing results have an average of  $>30$  million reads per sample. 26,799 annotated features can be identified with 64% features detectable without fragments per kilobase of transcript per million mapped reads (FPKM) cutoff; number of detectable features dropped to 18% when FPKM  $\geq 1$ . Sex can be discriminated using blood-based sex-specific gene set identified by the Genotype-Tissue Expression consortium. Here, we demonstrate the feasibility to acquire biologically-relevant gene expression from DBS using RNA-seq which provide a new avenue to investigate perinatal diseases in a high throughput manner.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Gene expression analysis captures a snapshot of cellular activity that reflects the response to genetic, environmental and epigenetic changes in a biological system. Determining phenotypes at the molecular level with differentially-expressed genes, using high-throughput technologies such as microarray and next-generation RNA-sequencing (RNA-seq), could help bridge the current gap between genotypes and phenotypes [1]. Neonatal dried blood spots (DBS) represent a unique biospecimen resource that can provide the “baseline” gene expression at birth for genotype-phenotype studies. In many developed countries, DBS are routinely collected on standard Guthrie cards for newborn screening to diagnose inborn errors of metabolism, hypothyroidism and other diseases [2]. Post-diagnostic gene expression analysis of DBS can potentially give insights into the cellular response of molecular pathological factors and thereby the etiology of a specific disease. For

example, it has been shown that microarray gene expression profiles from DBS can predict children with cerebral palsy [3].

Denmark has an all-comprising comprehensive system of registries which hold and collect information on social and health related life events for each resident linked through a unique person identifiable number (CPR number) [4–6]. Since the early 1980s, the Danish Neonatal Screening Biobank (DNSB) has been systematically collecting and storing surplus DBS from the neonatal screening program. While the primary purpose of DBS is to diagnose and treat congenital disorders, it can also be used for research purposes with appropriate approval [7, 8]. Several methods have been developed, primarily concerned with utilizing DNA from DBS [9–12]. In addition, we and several other groups have shown the feasibility of obtaining RNA and generating RNA gene expression microarray data from DBS for various studies [3, 13–19].

RNA microarray technologies are fast, robust and relatively inexpensive. However, they have inherent problems such as differing hybridization properties of probes, as well as predominant probe placements near the 3' end of transcripts. RNA-seq, a cutting-edge massively parallel next-generation sequencing technique, has the capacity to sequence each transcript on a single nucleotide basis. While above median transcripts are generally reliably detected by both RNA microarrays and RNA-seq, RNA-seq offers several significant advantages over microarray

Abbreviations: DBS, Dried blood spots; DNSB, Danish Neonatal Screening Biobank; GTEx, Genotype-Tissue Expression consortium.

\* Corresponding author at: Artillerivej 5 85#240, DK-2300 Copenhagen S, Denmark.

E-mail address: [jogr@ssi.dk](mailto:jogr@ssi.dk) (J. Bybjerg-Grauholm).

<sup>1</sup> These authors contributed equally to this work.

such as the capability to distinguish isoforms, call sequence variants, and provide improved accuracy for low-abundance transcripts [20–22]. However, the nature of total RNA, from blood in particular, does complicate sequencing approaches. Firstly, approximately 80% of total RNA from any tissue source is ribosomal RNA (rRNA). Thus, rRNA is usually reduced or removed prior to RNA-seq [23,24]. Secondly, in blood, 50–90% of mRNA transcripts are globin species [21,25]; globin may contribute to uninformative reads that are present in high abundance thus compromising RNA-seq analysis. Globin depletion can remove ~80% of globin species, improve correlation of technical replicates, and enable an additional 3500 RNA transcripts to be detected [21]. In neonatal DBS, globin is primarily derived from fetal hemoglobin (HbF), the predominant form of hemoglobin in the developing fetus which persists in newborn blood until gradually switching over to adult hemoglobin (HbA) ~ 6 months post-birth [26,27]. HbF could be depleted using the same methodology as HbA, but it remains untested whether sequence similarity is high enough to use the exact same method.

To date, RNA-seq for neonatal DBS has not been reported. Here, we conducted a proof-of-concept study to demonstrate the feasibility of obtaining robust and biologically-relevant gene expression data from DBS using RNA-seq.

## 2. Methods

### 2.1. DBS samples

This study does not constitute a health-related research project as defined by the Danish “Act on Research Ethics Review of Health Research Projects”. It is considered a developmental project for the Newborn Screening Program instead which does not require a separate approval from the Committees on Biomedical Research Ethics for the Capital Region of Denmark. In this study, we used a set of anonymized neonatal DBS samples that had been stripped of all information including the year of sampling. Sexes were determined using Sequenom's Sample ID panel (data not included). Sample names are PKU\_M\_0 to PKU\_M\_4 for five male DBS and PKU\_F\_0 to PKU\_F\_4 for five female DBS.

### 2.2. RNA isolation, library preparation and sequencing

Total RNA from two 3.2 mm DBS punches was extracted, purified and concentrated using Illustra RNAspin mini (GE Healthcare) and then subjected to DNase treatment to eliminate DNA contamination. Samples were concentrated using RNA Clean & Concentrator (Zymo Research). Sequencing libraries were prepared according to the standard protocol for the Stranded TruSeq RiboZero-Globin kit (Illumina) which includes Globin and rRNA depletion. RNA was fragmented for 4 min at 94 °C and each library was quantified using the KAPA library quantification kit (KAPA). Libraries were pooled before sequencing with a NextSeq-500 sequencer using NextSeq Control Software v1.3.0 (Illumina) on a high output flow cell (v1) with paired end reads of 76 bp (Illumina). Following sequencing, data was de-multiplexed and converted to FASTQ using bcl2fastq v2.17.1.14 (Illumina).

### 2.3. Data analysis

The FASTQ files were processed with the settings described by Trepnell [28] using Tophat v2.0.13 and Cufflinks v2.2.5. Filtering of genes and visualization of results was performed in cummerbund 2.8.2 [28] using the commands described by Trepnell [28] within the supplementary information for that paper. The reference genome and gene annotations used was UCSC HG19 prepackaged and downloaded via iGenomes (Illumina, downloaded June 2014).

## 3. Results

### 3.1. Library preparation and sequencing statistics

More than 325 million reads were generated from the 10 samples. The average reads per sample was 32.7 million  $\pm$  18.3 million, with 34.4 million  $\pm$  18.7 million for females and 31.1 million  $\pm$  20.0 million for males. The read length was 32–75 bp with an average of 74 bp, 91.5% of bases exceeded Q30. Detailed information of each sample's preparation and sequencing is shown in Table 1. Variation was observed in the dynamic range of transcript intensities, i.e. fragments per kilobase of transcript per million mapped reads (FPKM). Variance in FPKM was more profound across samples than between the sexes (Supplementary Fig. 1).

### 3.2. Transcript detection

To evaluate transcript detection rates, we limit the read alignment to redundant annotated features (transcripts and non-coding RNA). In total, our reference contained 26,799 features. An average of 64% of all features were detected (SD = 15%, range 32–81%) without FPKM cutoff while the number of detectable features dropped to 18% (SD = 7%, range 3–26%) when FPKM  $\geq$  1 was applied.

To test for potential outliers, dimensional reduction analysis with Multi-Dimensional Scaling (MDS) and Principle Component Analysis (PCA) were used. Both MDS and PCA exhibited two same potential outliers PKU\_M\_3 and PKU\_F\_1, with MDS showed a potential third outlier PKU\_M\_0 (Supplementary Fig. 2). We did not exclude the outliers from subsequent analysis of differential expression analysis based on sexes.

### 3.3. Differentiation of sex-specific gene expression

To verify whether DBS can reveal biologically-relevant gene expression using RNA-seq, we evaluated a gene set (19 sex chromosome-linked genes; 21 features) identified by the Genotype-Tissue Expression (GTEx) consortium [29] that are differentially-expressed between male and female in blood. Hierarchical clustering from FPKM showed complete separation of males from females (Fig. 1A). Gene expression heat map showed female samples expressed only X chromosome-linked genes (*XIST* and *TXLNG*) and none of Y chromosome-linked genes while male samples expressed >50% of Y chromosome-linked genes (Fig. 1B). FPKM values for each sex-specific gene are shown in Supplementary Table 1.

### 3.4. Effect of globin depletion

In this study, although the exact efficiency of globin depletion is unquantifiable, our results showed that globin did not represent the majority of transcripts. To access the effect of globin depletion on RNA-seq data, FPKM values of eight hemoglobin transcripts (*HBBP1*, *HBG1*, *HBG2*, *HBE1*, *HBZ*, *HBM*, *HBA1* and *HBQ1*) were used to create hierarchical clustering and heatmaps. PKU\_M\_1 has higher FPKM values for *HBG1* and *HBG2* when compared with the rest of the samples. Interestingly, by using the entire hemoglobin gene set in the analysis, female samples can be discriminated from male (Fig. 2A and B). However, when *HBA1* is removed from the analysis, sexes of the samples could not be differentiated (Fig. 2C and D). Analysis using the full set of transcripts separated the three previously identified outliers into another cluster (Fig. 2E). FPKM values for hemoglobin transcripts are shown in Supplementary Table 2.

## 4. Discussion

Our results show that we can perform RNA-seq on archived DBS samples, despite having some outliers in the dimensional reduction analysis. Sufficient data can be generated for hypothesis driven research

**Table 1**  
RNA input and RNA-seq data of 10 dB (5 females and 5 males).

Sample ID	Input RNA (ng)	Read length range (bp)	Average read length (bp)	Raw bases Q10+ (bp)	Raw bases Q20+ (bp)	Raw bases Q30+ (bp)	Read count	Total sequence generated (bp)
PKU_F_0	314	32–75	74.51	99.99	99.75	93.18	41,164,517	3,067,015,571
PKU_F_1	70	32–75	74.48	99.99	99.52	93.64	49,891,753	3,715,888,452
PKU_F_2	40	32–75	74.57	99.98	99.74	93.99	47,901,645	3,571,868,192
PKU_F_3	40	35–75	74.41	99.87	99.42	89.84	28,409,088	2,113,940,315
PKU_F_4	<0.2	35–75	63.65	76.11	51.85	44.73	4,520,783	287,745,978
PKU_M_0	60	32–75	74.48	99.99	99.68	93.28	45,422,192	3,382,887,870
PKU_M_1	500	32–75	74.51	99.99	99.81	94.51	19,988,616	1,489,344,449
PKU_M_2	1000	35–75	74.44	99.88	99.44	88.95	58,476,079	4,352,678,456
PKU_M_3	1000	35–75	74.54	99.96	99.48	89.81	20,580,446	1,533,986,335
PKU_M_4	1000	35–75	74.54	99.97	99.48	89.82	10,949,721	816,165,766

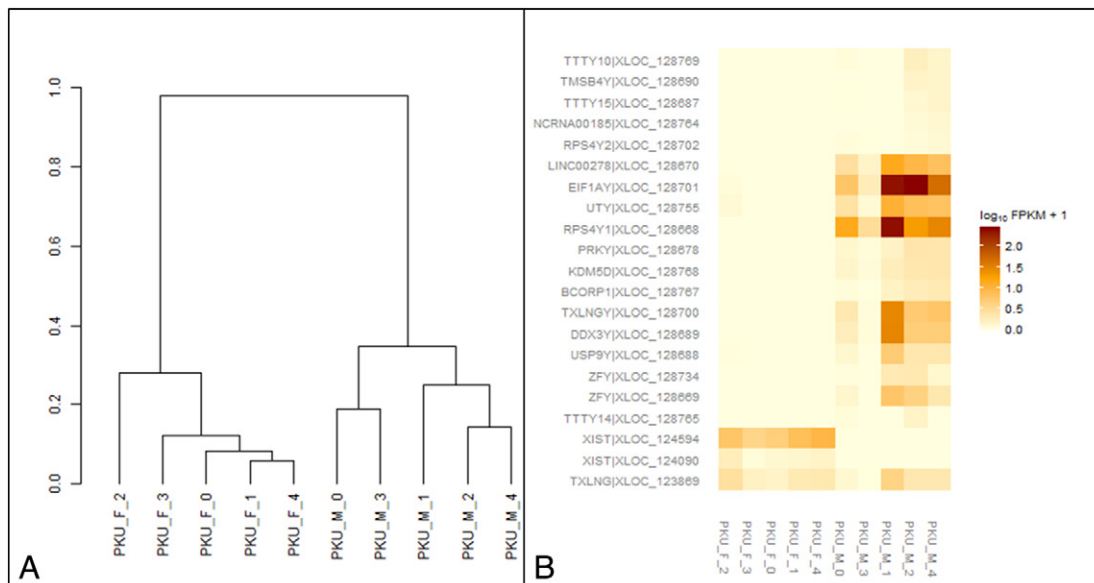
such as gender differentiation shown in this study (gender were correctly called 10 out of 10). Thus, it is feasible to perform RNA-seq on neonatal DBS for transcriptome study shortly after birth. While some biological questions can be answered with DNA sequencing, other inquiries such as pathway analysis for specific disease as demonstrated by Ho et al. in cerebral palsy shows gene expression studies n DBS can contribute to new knowledge [3].

RNA input for our library preparation was restricted to two 3.2 mm DBS punches for each sample. This imposed limitation is not an arbitrary decision as it reflects what is typically granted by the DNSB steering committee. The steering committee sets this limitation in order to ensure sample availability for the primary purpose of diagnostics, but also availability for future research projects. The amount of total RNA from two punches of 3.2 mm DBS ranges from <0.2 ng to 1 µg while the Ribo-Zero Globin reduction kit calls for 100 ng to 1 µg from either fragmented or intact total RNA input to enhance transcript coverage. Consequently, some of the preparations yielded less than the required library output before sample pooling. In cases with low library output, we used all available material and accepted a reduction in the expected read count. Thus, the RNA-seq read counts varied among individual samples. However, the read count difference is negligible when comparing group averages, as shown by 34 million and 31 million reads for females and males respectively.

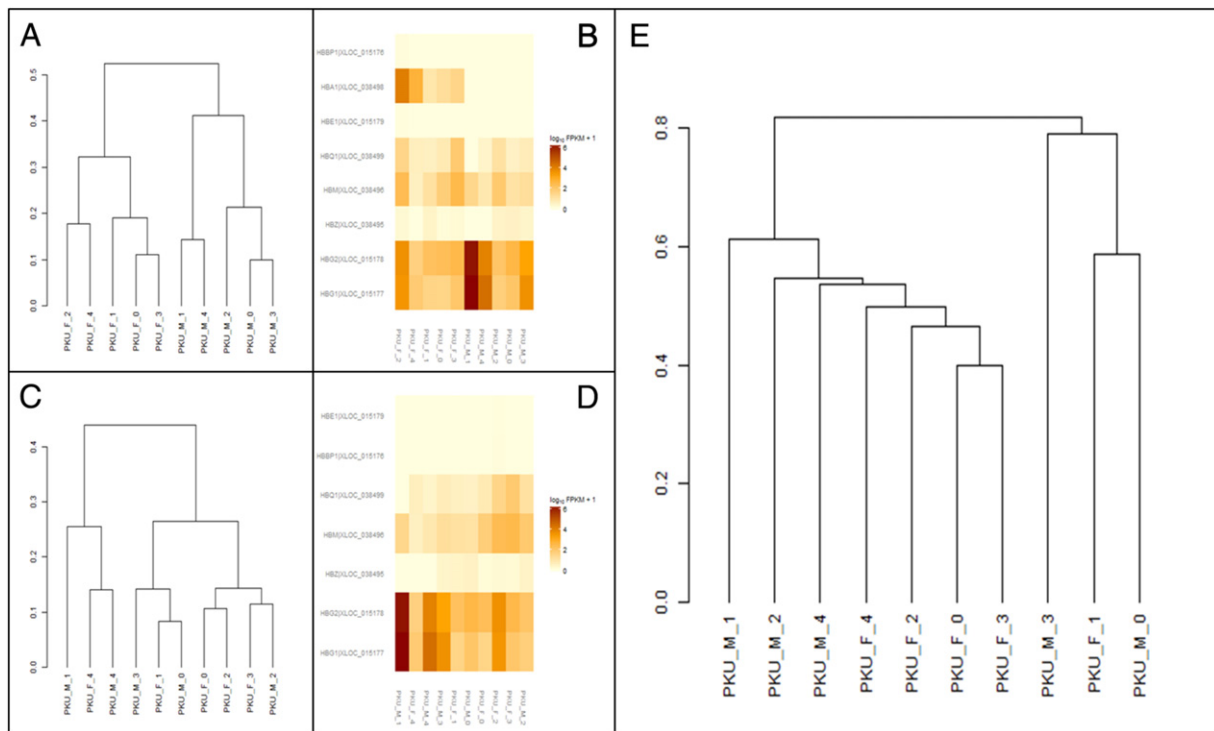
Here, we detected an average of 65% annotated features (genes and non-coding RNA) at an average reads count of 32.7 million in 10 DBS samples. With FPKM ≥ 1, the number of detectable features was reduced

to 18% compared with 33% in a previous study using globin reduction in human blood with ~30 M mapped reads [21]. A study on mouse Th2 cells reported ~70% detected genes at ~25 million reads when there is no FPKM cutoff [30]. This suggests that the number of annotated features depends on several factors such as read counts, sample types, and types of organism.

As HbA and HbF are gene complexes, we are unable to quantify their depletion directly. Thus, we evaluated the individual globin subunit transcripts within our dataset instead. While our experiment did not allow us to assess the effectiveness of the depletion directly, we can deduce the globin depletion effect by comparing our data with those in the public datasets. In GTEx, average FPKM values in blood for *HBA1* and *HBA2* are reported at 43,438 and 121,010 respectively. In our study, *HBA1* has an average FPKM of 1318 and *HBA2* is undetected (Supplementary Table 2). Thus, the majority of these globin species are been depleted in our study. Interestingly, hierarchical clustering of a globin gene enrichment data set showed two distinct clusters for females and males, driven by *HBA1* (Fig. 2A and B). When *HBA1* was removed from the analysis, hierarchical clustering and heat map showed a mixture of males and females (Fig. 2C and D). This observation is in concordance with a study which suggested that sex-associated factor may affects globin switching from HbF to HbA, showing HbA concentration differed significantly between female and male newborns [31]. A mixture of female and male clusters is also shown when using the entire gene set (unfiltered) (Fig. 2E). In addition, others have also shown that globin depletion does not introduce significant bias in biological and technical



**Fig. 1.** A) Hierarchical clustering of FPKM values based on differentially-expressed genes between males and females in blood genes identified by the GTEx consortium. B) Heatmap of FPKM values based on differentially-expressed genes between males and females in blood genes identified by the GTEx consortium.



**Fig. 2.** Hierarchical clustering (HC) and heat maps (HM) based on a gene set enrichment of hemoglobin species. A) HC of all the identified blood genes. B) HM of all the identified blood genes. C) HC of the identified blood genes, excluding *HBA1*. D) HM of all the identified blood genes, excluding *HBA1*. E) HC based on the entire gene set (unfiltered).

replicates [21]. This finding further support the feasibility of using DBS to detect reliable biologically-relevant gene expression.

Our experimental design was intended to be simple and hypothesis driven, since the inference we wish to draw upon in this proof-of-principle study is that if we can determine sex, then other biologically significant signals could also be detected. We chose to maintain this simplicity by strictly following the established RNA-seq data analysis pipeline published by Tranell et al. [28], thereby applying an established approach and avoid arbitrary decision bias of the findings. When enriching our dataset with sex-specific transcripts in blood determined by GTEx, we are able to distinguish females from males in the DBS samples, as presented by hierarchical clustering and heat map in Fig. 1. The separation of female from male cluster is close to 1, indicating a near complete segregation which agrees with the sex call from the Sequenom Sample ID panel (data not shown). However, not all sex-segregating transcripts are consistently expressed at detectable levels (Supplementary Table 1). The exact reason is unclear. One explanation can be the age difference between the GTEx cohort and DBS samples: DBS are taken from newborns while GTEx samples are taken from post-mortem donors aged 21–70 [32,33]. Despite some discordance with the GTEx data, there are a number of relevant candidate genes for sex determination in neonatal samples within the GTEx sex-specific set. From Figs. 1 and 2, as well as Supplementary Tables 1 and 2, we found the expression of *LINC00278*, *EIF1AY*, *UTY* and *RPS4Y1* unique in males, while *XIST* and *HBA1* are unique in females. These genes are suitable to determine sex in DBS.

In addition to laboratory research applications, DBS technology can be expanded to patients of any age as a standard blood collection to simplify clinical procedures. Depending on the purpose of clinical visitation, DBS can be drawn from a finger in the patient's own home to reduce or eliminate frequent visits to the clinic. It would also minimize the discomfort to a prick in the finger compared to a standard blood draw. DBS may also facilitate clinical diagnosis. For example, fetal hypoxia that leads to more than a million stillbirths worldwide each year can be diagnosed with the presence of fetal induced hypoxia but requires multiple testing at multiple times [34,35]. DBS can be collected at

multiple times with minimum stress to measure hypoxia-induced RNA and identify critically hypoxic fetuses in utero. There is also a potential to collect blood as DBS samples in poor-resourced clinical settings as this methodology greatly simplifies the logistics and reduce processing costs. It has been recommended to perform microarray differentially gene expression analysis from DBS stored at ambient temperature within 6 years while there is no gene expression difference between samples stored at 10 years apart at  $-20^{\circ}\text{C}$ . However, the effects of storage time and temperature on RNA-seq data obtained from DBS is currently unknown and warranted future investigation [16,36].

In conclusion, this proof-of-concept study demonstrated the feasibility to obtain reliable gene expression from DBS using RNA-seq. It is also feasible to use globin depletion protocols intended for adult hemoglobin for fetal globin depletion in DBS. Sex can be differentiated from DBS and a combination of six transcripts: *LINC00278*, *EIF1AY*, *UTY*, *RPS4Y1*, *XIST* and *HBA1* allowed high confidence of sex determinations. We can also determine novel biological changes such as differences in HbA levels between males and females. This newly established RNA-seq assay for DBS will enable high-throughput molecular epidemiology and diagnostic studies to better understand various perinatal diseases.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jymgmr.2016.12.004>.

#### Authors' contributions

J.B.-G. designed the study and supervised the laboratory work, analyzed and interpreted the data and drafted the manuscript.

C.M.H. analyzed and interpreted the data and drafted the manuscript.

S.K.K. consulted on the RNA-extractions and participated in interpretation of the results and writing of the manuscript.

M.L.J. involved in the study design and did the majority of the laboratory work.

C.S.H. and M.B.-H. participated in the handling of the samples and the interpretation of the results.

M.V.H., M.C. and D.M.H. initiated the study, participated in its design and the interpretation of results.

All authors critically revised and approved the manuscript.

### Competing financial interests

The authors declare no competing financial interests.

### Acknowledgements

The authors would like to thank and acknowledge our dear deceased friend and colleague Mads Vilhelm Hollegaard. We are but dwarfs perched on the shoulders of giants and thus see further than they do. In all our work with Mads, he helped us see the furthest horizons.

### References

- [1] Y.A. Kim, T.M. Przytycka, Bridging the gap between genotype and phenotype via network approaches, *Front. Genet.* 3 (2012) 227.
- [2] N.S. Green, K.A. Pass, Neonatal screening by DNA microarray: spots and chips, *Nat. Rev. Genet.* 6 (2) (2005) 147–151.
- [3] N.T. Ho, et al., Gene expression in archived newborn blood spots distinguishes infants who will later develop cerebral palsy from matched controls, *Pediatr. Res.* 73 (4 Pt 1) (2013) 450–456.
- [4] L.C. Thygesen, et al., Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving, *Scand. J. Public Health* 39 (7 Suppl) (2011) 12–16.
- [5] M. Schmidt, L. Pedersen, H.T. Sorensen, The Danish Civil Registration System as a tool in epidemiology, *Eur. J. Epidemiol.* 29 (8) (2014) 541–549.
- [6] C.B. Pedersen, et al., The Danish civil registration system. A cohort of eight million persons, *Dan. Med. Bull.* 53 (4) (2006) 441–449.
- [7] B. Norgaard-Pedersen, D.M. Hougaard, Storage policies and use of the Danish newborn screening biobank, *J. Inherit. Metab. Dis.* 30 (4) (2007) 530–536.
- [8] M. Hartlev, Genomic databases and biobanks in Denmark, *J. Law Med. Ethics* 43 (4) (2015) 743–753.
- [9] M.V. Hollegaard, et al., Genome-wide scans using archived neonatal dried blood spot samples, *BMC Genomics* 10 (2009) 297.
- [10] M.V. Hollegaard, et al., Archived neonatal dried blood spot samples can be used for accurate whole genome and exome-targeted next-generation sequencing, *Mol. Genet. Metab.* 110 (1–2) (2013) 65–72.
- [11] M.V. Hollegaard, et al., DNA methylome profiling using neonatal dried blood spot samples: a proof-of-principle study, *Mol. Genet. Metab.* 108 (4) (2013) 225–231.
- [12] M.V. Hollegaard, et al., Robustness of genome-wide scanning using archived dried blood spot samples as a DNA source, *BMC Genet.* 12 (2011) 58.
- [13] Y.H. Zhang, E.R. McCabe, RNA analysis from newborn screening dried blood specimens, *Hum. Genet.* 89 (3) (1992).
- [14] S.K. Khoo, et al., Acquiring genome-wide gene expression profiles in Guthrie card blood spots using microarrays, *Pathol. Int.* 61 (1) (2011).
- [15] F. Gauffin, et al., Quantitation of RNA decay in dried blood spots during 20 years of storage, *Clin. Chem. Lab. Med.* 47 (12) (2009) 1467–1469.
- [16] J. Grauholm, et al., Gene expression profiling of archived dried blood spot samples from the Danish Neonatal Screening Biobank, *Mol. Genet. Metab.* 116 (3) (2015) 119–124.
- [17] J.H. Resau, et al., Evaluation of sex-specific gene expression in archived dried blood spots (DBS), *Int. J. Mol. Sci.* 13 (8) (2012) 9599–9608.
- [18] J. Slaughter, et al., High correlations in gene expression between paired umbilical cord blood and neonatal blood of healthy newborns on Guthrie cards, *J. Matern. Fetal Neonatal Med.* 26 (18) (2013) 1765–1767.
- [19] C. Wei, et al., Comparison of frozen and unfrozen blood spots for gene expression studies, *J. Pediatr.* 164 (1) (2014) 189–191.e1.
- [20] C. Wang, et al., The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance, *Nat. Biotechnol.* 32 (9) (2014) 926–932.
- [21] Shin, H., et al., Variation in RNA-Seq transcriptome profiles of peripheral whole blood from healthy individuals with and without globin depletion, *PLoS One*, 2014. 9(3): p. e91041.
- [22] J.C. Marioni, et al., RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays, *Genome Res.* 18 (9) (2008).
- [23] O'Neil, D., H. Glowatz, and M. Schlumpberger, Ribosomal RNA Depletion for efficient use of RNA-seq capacity. *Curr. Protoc. Mol. Biol.*, 2013. Chapter 4: p. Unit 4 19.
- [24] W. Zhao, et al., Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling, *BMC Genomics* 15 (2014) 419.
- [25] A. Mastrokolias, et al., Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA, *BMC Genomics* 13 (2012) 28.
- [26] I. Akinsheye, et al., Fetal hemoglobin in sickle cell anemia, *Blood* 118 (1) (2011) 19–27.
- [27] K. Adachi, et al., Characterization of two types of fetal hemoglobin: alpha 2G gamma 2 and alpha 2A gamma 2, *Blood* 75 (10) (1990) 2070–2075.
- [28] C. Trapnell, et al., Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks, *Nat. Protoc.* 7 (3) (2012).
- [29] The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.*, 2013. 45(6): p. 580–5.
- [30] D. Hebenstreit, et al., RNA sequencing reveals two major classes of gene expression levels in metazoan cells, *Mol. Syst. Biol.* 7 (2011) 497.
- [31] F. Galacteros, M. Guillaud-Bataille, J. Feingold, Sex, gestational age, and weight dependency of adult hemoglobin concentration in normal newborns, *Blood* 78 (4) (1991) 1121–1124.
- [32] G.T. Consortium, The genotype-tissue expression (GTEx) project, *Nat. Genet.* 45 (6) (2013) 580–585.
- [33] J.C. Keen, H.M. Moore, The genotype-tissue expression (GTEx) project: linking clinical data with molecular analysis to advance personalized medicine, *J. Pers. Med.* 5 (1) (2015) 22–29.
- [34] J.E. Lawn, et al., Stillbirths: where? when? why? How to make the data count? *Lancet* 377 (9775) (2011) 1448–1463.
- [35] C.L. Whitehead, S. Tong, Measuring hypoxia-induced RNA in maternal blood: a new way to identify critically hypoxic fetuses in utero? *Expert. Rev. Mol. Diagn.* 14 (5) (2014) 509–511.
- [36] N.T. Ho, et al., Effect of storage time on gene expression data acquired from unfrozen archived newborn blood spots, *Mol. Genet. Metab.* 119 (3) (2016) 207–213.