Biochemical and Biophysical Research Communications xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Pairwise comparisons of ten porcine tissues identify differential transcriptional regulation at the gene, isoform, promoter and transcription start site level ☆

Leila Farajzadeh, Henrik Hornshøj [1], Jamal Momeni, Bo Thomsen, Knud Larsen, Jakob Hedegaard [1], Christian Bendixen, Lone Bruhn Madsen *

Department of Molecular Biology and Genetics, Faculty of Sciences and Technology, Aarhus University, DK-8830 Tjele, Denmark

ARTICLE INFO

Article history:
Received 15 July 2013
Available online xxxx

Keywords:
Gene expression
Pig
Promoter
RNA-seq
Splicing

ABSTRACT

The transcriptome is the absolute set of transcripts in a tissue or cell at the time of sampling. In this study RNA-Seq is employed to enable the differential analysis of the transcriptome profile for ten porcine tissues in order to evaluate differences between the tissues at the gene and isoform expression level, together with an analysis of variation in transcription start sites, promoter usage, and splicing.

Totally, 223 million RNA fragments were sequenced leading to the identification of 59,930 transcribed gene locations and 290,936 transcript variants using Cufflinks with similarity to approximately 13,899 annotated human genes. Pairwise analysis of tissues for differential expression at the gene level showed that the smallest differences were between tissues originating from the porcine brain. Interestingly, the relative level of differential expression at the isoform level did generally not vary between tissue contrasts. Furthermore, analysis of differential promoter usage between tissues, revealed a proportionally higher variation between cerebellum (CBE) versus frontal cortex and cerebellum versus hypothalamus (HYP) than in the remaining comparisons. In addition, the comparison of differential transcription start sites showed that the number of these sites is generally increased in comparisons including hypothalamus in contrast to other pairwise assessments.

A comprehensive analysis of one of the tissue contrasts, i.e. cerebellum versus heart for differential variation at the gene, isoform, and transcription start site (TSS), and promoter level showed that several of the genes were differed at all four levels. Interestingly, these genes were mainly annotated to the "electron transport chain" and neuronal differentiation, emphasizing that "tissue important" genes are regulated at several levels. Furthermore, our analysis shows that the "across tissue approach" has a promising potential when screening for possible explanations for variations, such as those observed at the gene expression levels.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

1. Introduction

Genomic and transcriptomic data are accumulating at an exceptional rate allowing for the elucidation on how particular functions embodied by specific tissues appear from a universal set of molecular instructions. The understanding of this link is especially challenging in the mammalian brain, where the complexity derives from the abundance and diversity of cell types from which it is comprised [19]. For various tissues including neuronal cell types deep sequencing of RNAs (RNA-seq) has recently been applied to quantify gene and alternative isoform expression levels [13,27,29]. In RNA-seq experiments all RNAs of a particular sample are randomly fragmented, reverse transcribed, ligated to adapters, and subsequently the fragments are sequenced and levels of gene expression are derived from the number of sequence reads originating from each gene. Furthermore, RNA-seq also allows for the detection of transcription start sites and splice variants, the latter being highly important regarding mammalian brain complexity [15].

Furthermore, it is noteworthy that both genetically engineered pig models and pigs with naturally occurring mutations often more closely recapitulate human disease symptoms than the smaller

☆ This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author. Address: Department of Molecular Biology and Genetics, Aarhus University, Blichers Allé 20, Postboks 50, 8830 Tjele, Denmark. Fax: +45 87154994.

E-mail address: LoneB.Madsen@agrsci.dk (L.B. Madsen).

[1] Present address: Department of Molecular Medicine (MOMA), Aarhus University Hospital, Aarhus N DK-8200, Denmark.

0006-291X/$ - see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.07.074

models [20,23]. Moreover the size of pig genome, gene location and length of coding and non-coding regions resemble that of humans, which for instance has been shown in the characterization of genes important in neurological diseases such as Dystonia [10] and neurodegenerative diseases as Parkinson's and Alzheimer's disease [14,16]. For the past two decades pigs have been used in biomedical research with increasing frequency due to their close resemblance to humans regarding anatomy and physiology [26]. Furthermore, characterization of trinucleotide repeats in genes crucial for the development of neurological diseases associated with expansion of trinucleotide repeat elements, such as Huntington's disease in pigs, shows that these repeats have the potential to expand in pigs, making this species a potential natural model for these diseases [17].

In order to detect abnormal transcription profiles in disease state, it is essential to obtain knowledge of transcription profiles in various tissues including brain regions under normal, healthy conditions. In the current study we have employed RNA-seq on ten different porcine tissues including four brain tissues from two boars to gain an understanding of the variations in transcriptional profiles for these tissues consisting of occipital cortex (OCC), frontal cortex (FCO), hypothalamus and cerebellum along with such diverse tissues as heart, spleen, liver, kidney, lung and *musculus longissimus dorsi*. This has enabled us to perform comparative gene expression analysis of brain regions versus non-brain tissues along with inter-brain tissue comparisons. Hence, we have tested for differentially expressed genes and isoforms, differential splicing, TSS, and differential promoter usage between all ten porcine tissues. Furthermore, our analysis shows that the "across tissue approach" has a promising potential when screening for possible molecular explanations for variations in for instance gene expression levels.

2. Materials and methods

2.1. Animals and sample preparation

Ten tissues from two unrelated one year old Landrace boars were included in the study. Hence, total RNA was extracted from heart, spleen, liver, kidney, lung, *musculus longissimus dorsi*, occipital cortex, hypothalamus, frontal cortex, and cerebellum employing the mirVana™ RNA extraction kit (Ambion) according to manufactures protocol, yielding a total of 20 samples. RNA integrity of the individual RNA samples was assessed on a 2% agarose gel.

2.2. Library preparation and sequencing

Library preparation was performed using the mRNA-seq library prep kit from Illumina according to manufacturer's protocol. Briefly, 10 μg total RNA of each sample was employed in polyA mRNA selection using magnetic beads, followed by thermal fragmentation, resulting in a total of 10 samples for each of the two animals. Subsequently, the fragmented mRNA was reverse transcribed using reverse transcriptase (SuperScript II) and random primers followed by second strand cDNA synthesis using DNA polymerase I. The cDNA was end repaired using T4 DNA polymerase and Klenow DNA polymerase and size selected on a low melting 2% agarose gel and fragments corresponding to sizes of 200 nucleotides were excised from the gel and DNA was recovered applying the QIAquick gel extraction kit (Qiagen). Finally, in order to amplify the libraries, 15 cycles of PCR were employed using primers from the multiplexing sample preparation Oligonucleotide kit (Illumina). The amplified libraries were purified using the QIAquick PCR purification kit (QIAGEN) and the concentration of each library was determined using the Qubit flurometer (Invitrogen). The purified libraries were denatured and diluted to a concentration of 10 nM, and sequenced (50 bp single-end) as multiplex of ten samples per lane in a total of two lanes on a Genome Analyzer (Illumina). Output was transferred to the pipeline computer (Illumina Pipeline version 1.4), quality filtered, and analyzed using the Genome Analyzer Pipeline Software generating the raw fastq files. The fastq files and the processed data have been submitted to ArrayExpress with the accession number E-MTAB-1405.

2.3. Mapping and assembly of fragments

The quality filtered fastq files were mapped to the *Sus scrofa* reference genome build 10.2 [9] employing TopHat version 1.0.12 [24] using default settings, implying that reads were split in segments of 25 bp and mapped independently to the reference genome. Next, Cufflinks version 0.8.0 [25] was applied to assemble the aligned reads into clusters of overlapping reads, then building a graph which represents all possible isoforms, followed by traversing the created graph to assemble isoforms by finding the minimum set of transcripts that explain the intron junctions within the sequenced reads, allowing for the detection of new isoforms and transcripts. Additionally, the normalized RNA-seq fragment counts were used in the Cufflinks program to estimate the relative abundance of each transcript and are reported in fragments per kilobase of exon per million fragments mapped (FPKM). Transcripts were annotated to the human Refseq database downloaded at July 6th 2012.

2.4. Differential expression

Test for differential expression at the gene and isoform level was accomplished using Cuffdiff on applying all XLOC IDs established by Cufflinks. Furthermore, test for differential splicing, promoter usage and TSS was also accomplished using the Cuffdiff program. The tests were performed pairwise between all ten tissues yielding a total of 45 tests per condition and all tests were performed using default settings in Cuffdiff. Type I errors were evaluated using the Benjamini–Hochberg (1995) correction for multiple testing [3] and a significance level of 0.05. Gene-annotation enrichment analysis and functional annotation clustering were performed using transcripts annotated to the human Refseq database.

2.5. Visualization of spliced alignments

In silico visualization of splice data was accomplished using a local copy of the Integrative Genomics Viewer (IGV) software available at http://www.broadinstitute.org/igv/ [22] and establishment of Venn diagrams was completed in Venny [21].

3. Results and Discussion

It has previously been shown that the pig in terms of genetics, anatomy, and physiology mimics humans extremely well, which is emphasized by the increased use of pigs in biomedical research [26]. In this study massive parallel sequencing was employed to gain a comprehensive understanding of the transcriptome profile including variation in isoforms, TSS and promoter usage between various porcine tissues including hypothalamus, frontal cortex, occipital cortex, and cerebellum.

UNCORRECTED PROOF

Please cite this article in press as: L. Farajzadeh et al., Pairwise comparisons of ten porcine tissues identify differential transcriptional regulation at the gene, isoform, promoter and transcription start site level, Biochem. Biophys. Res. Commun. (2013), http://dx.doi.org/10.1016/j.bbrc.2013.07.074

Download English Version:

https://daneshyari.com/en/article/10758087

Download Persian Version:

https://daneshyari.com/article/10758087

Daneshyari.com