Nanopore sequencing unveils the complexity of the cold-activated murine brown adipose tissue transcriptome

5   Christoph Andreas Engelhard[1], Sajjad Khani[2,3], Sophia Derdak[4], Martin Bilban[5,#] and Jan-Wilhelm Kornfeld[1,#,*]

[1] Department for Biochemistry and Molecular Biology (BMB), University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark. [2] Max Planck Institute for Metabolism Research, Gleueler Strasse 50, 50931 Cologne, Germany. [3] Cologne Excellence Cluster on Cellular Stress

10   Responses in Ageing-Associated Diseases (CECAD), University of Cologne, Cologne, Germany. [4] Core Facilities, Medical University of Vienna, Lazarettgasse 14, 1090 Vienna, Austria. [5] Department of Laboratory Medicine & Core Facilities, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria. [#] These authors contributed equally. [*] Corresponding author. e-mail: janwilhelmkornfeld@bmb.sdu.dk and martin.bilban@meduniwien.ac.at

1.   Abstract

15   Alternative transcription increases transcriptome complexity by expression of multiple transcripts per gene and thus fine tunes cellular identity and function. Annotation and quantification of transcripts at complex loci using short-read sequencing is non-trivial. Recent long-read sequencing methods such as those from Oxford Nanopore Technologies

20   (ONT) and Pacific Biosciences aim at overcoming these problems by sequencing full length transcripts. Activation of BAT thermogenesis involves major transcriptomic remodelling and positively affects metabolism via increased energy expenditure and endocrine factors. Here we comprehensively benchmark features of ONT long-read sequencing protocols compared to Illumina shortread sequencing assessing alignment

25   characteristics, gene and transcript detection and quantification, differential gene and transcript expression, transcriptome reannotation and differential transcript usage (DTU). We find that ONT sequencing is superior to Illumina for transcriptome reassembly and reduces the risk of false-positive events due to the ability to unambiguously map reads to transcripts, at the expense of statistical power for calling differentially expressed features.

30   We identified novel isoforms of genes undergoing DTU in cold-activated BAT including

Cars2, Adtrp, Acsl5, Scp2, Aldoa and Pde4d, validated by RT-qPCR. Finally, we provide a reannotation of the murine iBAT transcriptome as a valuable resource for researchers interested in the molecular biology underlying the regulation of BAT.

2. Key words

35 Brown adipose tissue; metabolism, transcriptional regulation, alternative splicing, full length RNA-Seq

3. Abbreviations

BAT   brown adipose tissue

cDNA   complementary DNA

40 DTU   differential transcript usage

iBAT   interscapular brown adipose tissue

lncRNA   long non-coding RNA

mRNA   messenger RNA

ONT   Oxford Nanopore Technologies

45 PCR   polymerase chain reaction

qPCR   quantitative polymerase chain reaction

RNA   ribonucleic acid

TSS   transcription start site

4. Introduction

50 Alternative transcription (AT) is a post-transcriptional process in which multiple transcripts arise from a single gene locus by using alternative transcription start sites, altered polyadenylation sites and alternative splicing, thereby increasing the transcriptomic and translatomic complexity in a cell (de Klerk & 't Hoen, 2015). Alternative transcription is estimated to occur within 92% to 94% of human genes, substantially expanding the

55 catalogue of co-expressed mRNAs (Pan et al., 2008; E. T. Wang et al., 2008). In line, sequencing of ribosome attached translated mRNAs (Ribo-seq) and proteomics studies confirmed that many RNA species produced by AT are translated and contribute to increased proteome diversity (Floor & Doudna, 2016; Lau et al., 2019; Rodriguez et al., 2020; Weatheritt et al., 2016). Interestingly, AT is tissue specific (Rodriguez et al., 2020;

60   Xu et al., 2002) or marks specific cellular states (Fiszbein & Kornblihtt, 2017; Robinson et al., 2021), indicating a seminal role for AT in regulation of cellular identity and function.

Adipose tissue depots can be broadly classified into brown and white depots: While white adipocytes mainly function to store energy as triglycerides in large unilocular lipid droplets and coordinate energy metabolism by secretion of endocrine factors, brown adipocytes are

65   densely packed with mitochondria and morphologically present multiple small lipid droplets (Rosen & Spiegelman, 2014). Upon sympathetic nervous system activation e.g., upon cold stimulus, brown adipocytes upregulate lipolysis, where the ensuring free fatty acids activate Uncoupling Protein 1 (UCP1) and generate heat by increasing the uncoupling of oxidative respiration from ATP generation. Additionally, sympathetic activation of brown

70   adipocytes (BA) residing in brown fat and in inguinal white adipose tissue (so-called brown-in-white or 'brite' adipocytes) leads to profound changes in gene expression (Cannon & Nedergaard, 2004). Recent evidence suggests, that not only differential gene expression, but also the regulation of differential transcript usage (DTU) by RNA-binding proteins i.e., changes in the relative abundance of transcripts originating from one gene, is

75   crucial for e.g., the regulation of adipocyte thermogenesis (Vernia et al., 2016). Transcriptomic studies have shown that DTU is required for adipogenesis, the process of differentiation of preadipocytes into mature adipocytes (Fiszbein & Kornblihtt, 2017; Yi et al., 2020). Moreover, AT events in key brown adipocyte genes such as the transcription factors *Pparg* and *Prdm16* have been reported to play a role in the control of brown

80   adipocyte function (Chi & Lin, 2018; D. Li et al., 2016).

However, most studies focussing on AT and DTU so far have used Illumina short-read sequencing. Short-read sequencing inherently underperforms in relation to assembling transcripts, since the single reads only span a fraction of a transcript, and therefore requires complex computational post processing for transcriptome reassembly. This poses

85   a conceptual problem: If two individual AT events in one gene occur too far away from each other to be spanned by a single short read, it is challenging to unambiguously decide if both AT events happen (i) in conjunction, (ii) arise independently from each (iii) or are mutually exclusive (Byrne et al., 2019). Short-read sequencing also suffers in respect to transcript level quantification required for analysis of DTU, as only reads mapping to parts

90   of a gene unique to a single transcript can be unambiguously assigned to a transcript, while all others must be assigned based on statistical models (Bray et al., 2016; Patro et al., 2017). Long-read sequencing methods such as those developed by Pacific Biosciences (Eid et al., 2009; Sharon et al., 2013) and Oxford Nanopore Technologies (ONT; Garalde et al., 2018) generate full-length isoform reads that mitigate these

95   limitations, allowing for simple transcriptome reannotation and unambiguous read assignment (Stark et al., 2019). Importantly, thousands of novel transcripts across a large collection of different human tissues have recently been revealed using long-read sequencing with ONT, enabling an understanding of functionally distinct protein isoforms that different transcripts can give rise to (Glinos et al., 2022). Reference databases like

100   GENCODE are based on a limited number of tissue transcriptomes (Robinson et al., 2021; Xu et al., 2002). Since alternative transcription is tissue and cell state specific, it is of high biological interest to reannotate transcriptomes in cell types such as brown adipocytes, which are not represented in reference annotations, in order to identify and quantify tissue specific transcript isoforms. Long-read sequencing methods like ONT sequencing on the

105   other hand, suffer from lower throughput and lower base calling accuracy, resulting in failure to detect lowly expressed isoforms and fuzzy splice junction annotation (Stark et al., 2019). Accordingly, algorithms that combine short and long reads for improved transcriptome reassembly, either *de novo* or using a reference genome, have been developed (Fu et al., 2018; Kovaka et al., 2019; Tang et al., 2020) and consortia such as

110   GENCODE have started incorporating long-read sequencing in their reference transcriptome annotation pipelines (Frankish et al., 2021).

Here, we have compared three different library preparation methods using the ONT platform and assessed their ability for transcript detection, quantitation and differential expression calling in addition to performing transcriptome reassembly and analysis of

115   differential transcript usage. Using RNA isolated from murine iBAT, we identified cold induced isoform switches in genes regulating thermogenic β3 adrenergic receptor (AR) signalling at multiple levels including regulation of cAMP levels (*Pde4d*) and receptor signalling (*Adtrp*), lipid metabolism/signalling (*Scp2, Mlixpl*) and protein sorting (*Ergic1*). Finally, using FLAIR/ChIP Seq we identified a novel alternative transcription start site in

120   the in the mitochondrial respiration regulating protein *cysteinyl-tRNA synthetase gene (Cars2)* and validated coding potential for an alternative (shorter) transcript (*Cars2*-AT) using the *Coding-Potential Assessment* (CPA) Tool and determined functional domain structure using pfam. As an example which demonstrates the potential of the ONT long read iBAT transcriptome reannotation reported here, we show that sgRNAs targeting the

125   Cars2-AT promoter, are efficient in inducing the expression of Cars2-AT in brown adipocytes *in vitro*. Thus, we provide a reannotation of the murine iBAT transcriptome which can be a valuable resource for researchers interested in iBAT biology by facilitating them to target the relevant isoforms of a gene in study and detect novel DTU events in

cold activated murine iBAT, demonstrating the contribution of AT in the regulation of brown adipocyte activity.

5.   Results

Comparison of Nanopore-based approaches for transcript resolution of cold-activated BAT

We isolated RNA from iBAT of 20-week old, male C57BL/6N mice cold treated for 24 h at 4 °C or housed at room temperature (n = 3). To evaluate different library preparation methods for the ONT sequencing platform, we prepared libraries using *(i)* direct cDNA sequencing, which is PCR-free and avoids bias introduced by the amplification (Chen et al., 2021; Stark et al., 2019) and *(ii)* TeloPrime sequencing, which uses a 5' cap specific template switching oligo to enrich for full-length RNA but requires PCR amplification. All samples were multiplexed, and library pools sequenced on two separate flow cells per library preparation method on a ONT GridION to assess the variability in performance of the flow cells. Additionally, the samples were pooled within the respective treatment group and *(iii)* sequenced following ONT's direct RNA protocol on one flow cell each. Finally, *(iv)* we performed strand specific, paired-end short-read sequencing following Illumina's TruSeq protocol as reference (Fig. 1A). To the best of our knowledge, this analysis represents the most comprehensive characterization of full-length transcripts and transcript diversity to date in the murine BAT depot, both at basal levels and upon cold activation. Low quality ONT reads with a minimum average Phred Quality Score below 7 (20 % base call accuracy) were removed, leaving $13.5×10^6$ reads from TeloPrime sequencing, $12,7×10^6$ reads from direct cDNA sequencing, and $3,0×10^6$ from the direct RNA sequencinq (Tab. 1). Interestingly, we observed a high variability in the number of high-quality reads produced per flowcell, varying by 41 % for direct cDNA-Seq and 16 % for TeloPrime-Seq, even though the same libraries were used, and flow cells run in parallel (Fig. 1B). Similar variability in MinION flow cell performance has been noticed in other studies (Ip et al., 2015; Oikonomopoulos et al., 2016; Sessegolo et al., 2019). Noteworthy, a striking difference in high quality read numbers was seen between the samples harvested from mice housed at room temperature compared to those from cold treated animals in TeloPrime-Seq. Read length distributions were similar between samples and flow cells within one library preparation method (Fig. 1C). In good agreement with other reports (Udaondo et al., 2021), average read lengths were similar for direct RNA-Seq (1,033 nt) and direct cDNA-Seq (1,141 nt) but longer for TeloPrime (1,326 nt). The read

length distribution of the TeloPrime method however was multimodal, while the other distributions were unimodal. Average read quality was similar between samples, but interestingly depended on the flow cell used (Fig. 1D), in agreement with the number of high-quality reads received. Thus, our comparisons show that the TeloPrime protocol enriches for longer RNA molecules compared to other long-read protocols, but also reveal substantial technical variation between flow cells.
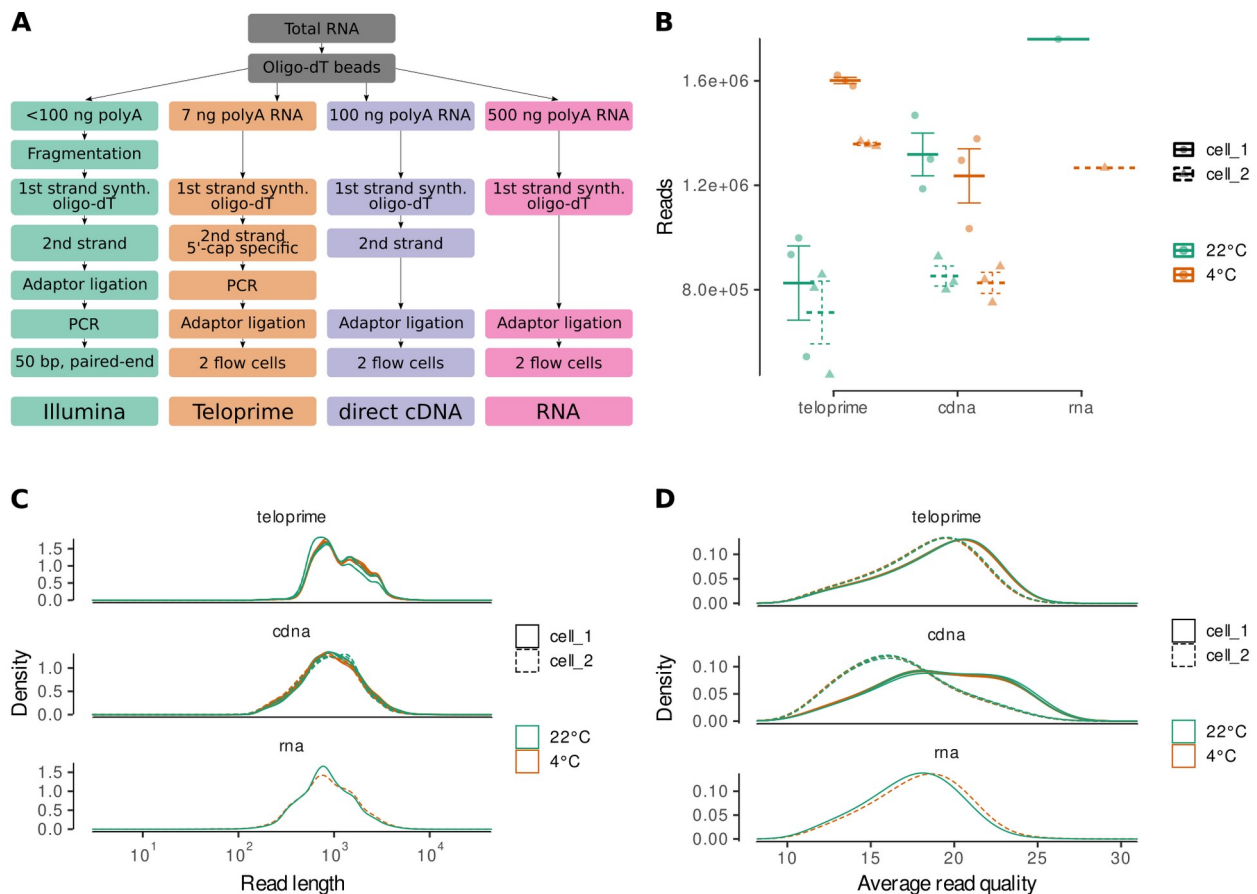

Fig. 1. Characterisation of ONT reads. **A** Experimental design. RNA was either sequenced by the Illumina short-read platform or by three different ONT long-read library preparation protocols: TeloPrime, direct cDNA sequencing (cdna) or direct RNA sequencing (rna). See results and methods for details. **B;C;D** Total number of reads (B), read length distribution (C) and read quality distribution (D) by ONT sequencing method, flow cell and housing temperature.

ONT teloprime improves coverage of full-length transcripts

We next aligned the quality filtered long reads to the murine genome and transcriptome using minimap2. Overall, alignment rates were high and independent of whether the alignment was performed using the transcriptome or genome as reference, ranging from 93 % for the direct cDNA-Seq method to > 99 % for TeloPrime (Tab. 1), emphasizing one of the main advantages for long-read RNA-seq. In line with other reports, we noted accompanying supplementary alignments i.e., reads not mapping linearly to the reference (Tab. 1; Fig. 2A; Fig. S1A-C). Reads with supplementary alignments were most common

in direct cDNA-Seq and these were longer (Fig. 2C; Fig. S1E) and showed larger unaligned parts (Fig. 2D; Fig. S1F) then reads from the other methods. Supplementary alignments can arise from reads mapping to different chromosomes, indicative for chromosomal rearrangements (Klever et al., 2020; Leung et al., 2021). However, in direct cDNA-Seq, supplementary alignments mostly mapped anti-sense to the same transcript as the primary alignment (Fig. 2B; Fig. S1D), indicating that the second strand of the cDNA was sequenced subsequently to the first strand. Next, we compared the ability of the different long-read sequencing methods to cover full transcripts. Comparison of the read length distribution of the different methods to the hypothetical distribution of transcript lengths as inferred from Illumina-Seq transcript abundances revealed that the direct cDNA-Seq method generated a read length distribution shifted towards shorter reads (Fig. S1G). This was less prominent for the TeloPrime and direct RNA-Seq methods. Investigation of gene body coverages revealed that Illumina short-read sequencing mostly covers the middle part of transcripts, with reduced coverage at the 3′ and 5' of the gene body in comparison to ONT sequencing, as demonstrated previously (Fig. 2E; Leshkowitz et al., 2022; Soneson et al., 2019; Wright et al., 2022). All long-read libraries showed decreasing coverage from the 3' end of transcripts towards the 5' end. Of note, this decrease was markedly reduced in TeloPrime-Seq, which is meant to enrich for full length transcripts. To assess the fraction of transcripts covered by reads and the proportion that represent full-length transcripts, a coverage fraction was calculated. We defined coverage fraction as the observed transcript length (alignment length) divided by the original known transcript length. We observed that both the coverage fraction as well as the fraction of full length reads markedly decreased with transcript length, in line with previous reports (Soneson et al. 2019; Fig. 2F; Fig. S1H). This could have been caused by RNA degradation during library preparation protocols or software artefacts during the base-calling process (Kovaka et al., 2019; Sessegolo et al., 2019). Overall, coverage was higher in TeloPrime-Seq as reported (Sessegolo et al., 2019). Direct cDNA-Seq showed both the lowest coverage and the smallest fraction of full length reads. As expected, the coverage of supplementary mappings was lower compared to primary alignments, although full-length supplementary alignments were present in all three ONT methods. Thus, we demonstrate that the TeloPrime protocol enriches for longer, potentially full length, transcripts compared to other two full-length sequencing protocols due to its dedicated library preparation method.
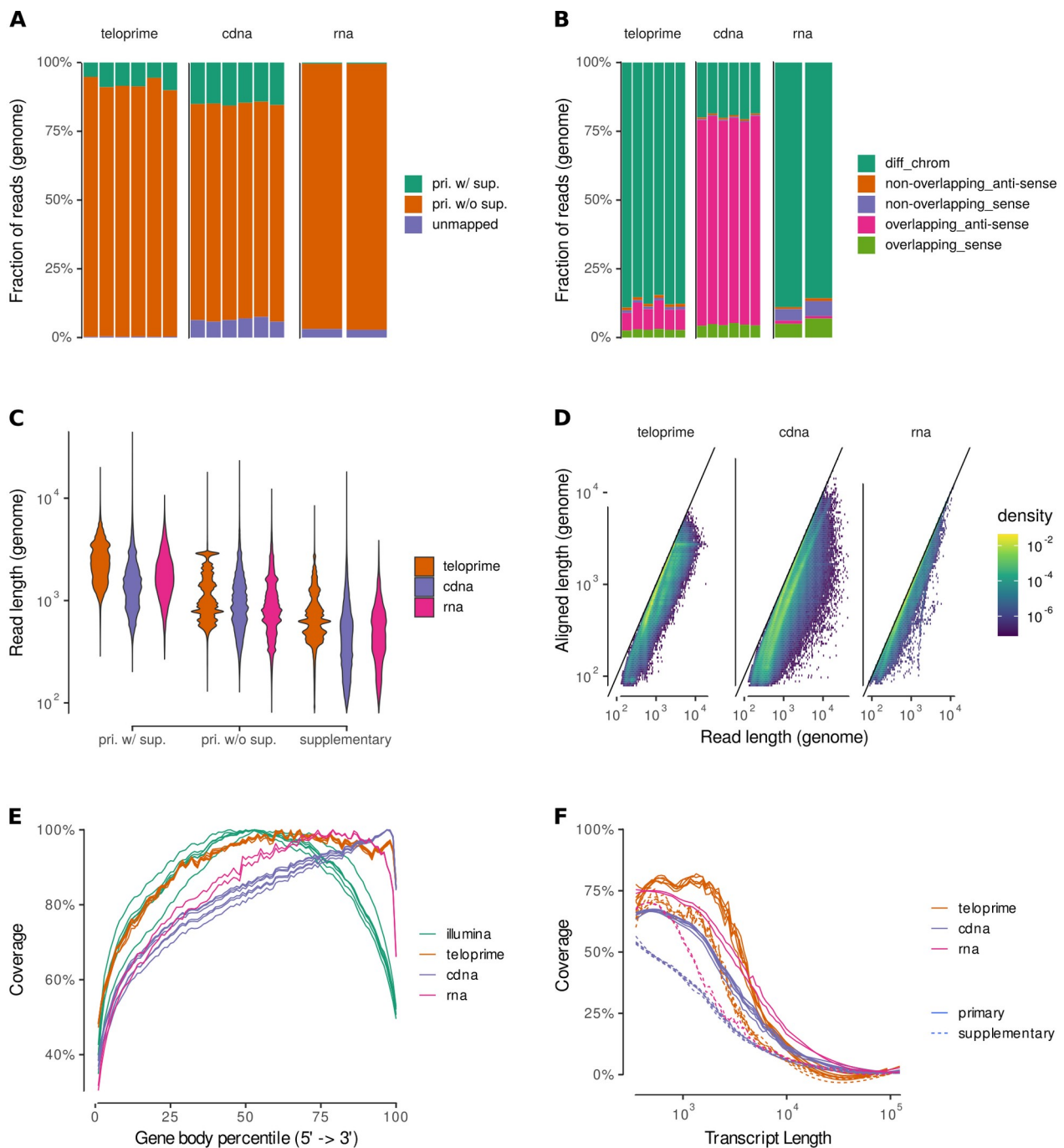
Fig. 2. Characterisation of read alignments. **A** Fraction of reads classified by whether the primary alignment against the genome has at least one supplementary alignment attached to it. **B** Fraction of supplementary alignments against the genome stratified by their relation to the corresponding primary alignment. **C** Read length distribution for reads aligned to the genome. **D** Aligned length vs read length for primary alignments to the genome. **E** Percentile wise coverage of the gene body based on primary genome alignments.
F Smoothed average transcript coverage of alignments mapped to the transcriptome for transcripts > 350 nt.

Tab. 1. Alignment characteristics. Number of total reads (Q > 7), alignment rates, rate of reads with supplementary alignments and number of supplementary alignments for the three long-read sequencing methods, stratified by mapping to the genome (GRCm38.p6) or transcriptome (GENCODE M22).

Method	Total reads	Reference	Alignm. rate	Prim w/ sup	Sup alignm.
teloprime	13495575	M22	99.6%	7.3%	1109736
		GRCm38.p6	99.7%	8.9%	1386229
cdna	12700471	M22	93.5%	15.0%	2031468
		GRCm38.p6	92.8%	22.5%	3247267
rna	3026502	M22	97.0%	0.4%	13385
		GRCm38.p6	96.2%	9.2%	311873

ONT direct RNA and cDNA protocols show less bias for gene/transcript detection compared with TeloPrime

We next compared the ability of the different sequencing methods to detect expressed genes and transcripts based on the reference annotation. In agreement with previous reports (Sessegolo et al., 2019), direct cDNA and RNA-Seq detected a similar number of features as short-read sequencing at any given sequencing depth far outperforming TeloPrime-Seq (Fig. 3A and B). While differences in sequencing depth explain the large sets of genes detected by Illumina-Seq or by direct cDNA-Seq and Illumina-Seq alone, we also observed 1447 genes only in direct RNA-Seq, indicating technical biases of the different methods. Similarly, most transcripts were observed in direct cDNA and Illumina-Seq. However, the share of transcripts only detected using one but not the other protocol was even more pronounced then on gene level, indicating differences in the transcript identification potential of the different technologies (Fig. 3C). To detect the cause of these differences, we stratified the transcript detection rates by transcript length and transcript biotype (Fig. 3D). As reported, detection rates for Illumina, direct cDNA-Seq and direct RNA-Seq increased with transcript size (Soneson et al., 2019). TeloPrime-Seq detection rates on the other hand were highest for transcripts ranging from 1000 nt to 3000 nt. Detection rates of the direct cDNA-Seq method reached the detection rates of short-read sequencing for coding genes longer than 5000 nt but not for long noncoding RNAs (lncRNAs). Genes and transcripts detected by either long- or short-read sequencing alone were enriched for noncoding RNA as compared to those detected by both sequencing types (Fig. S2A,B). Since coding genes are generally higher expressed than other classes of RNA (Cabili et al., 2011), we assessed a potential effect of expression level on the gene and transcript detection rates by the ONT methods (Fig. 3E; Fig. S2C). Both genes and transcripts detected by short-read sequencing and long-read sequencing showed on average a higher expression measured by Illumina-Seq compared to features detected by short-read sequencing only. However, there were also highly expressed genes and

transcripts that were not detected by the ONT sequencing protocols (Fig. S2D;E). Interestingly, there were also features with high expression in the ONT datasets that were 240 not detected by short-read sequencing, more prominently for transcripts compared to genes. Thus, direct cDNA-Seq and direct RNA-Seq had comparable gene and transcript detection rates, which were proportional to gene and transcript length, whilst TeloPrime yielded lower detection rates with a non-linear relationship to transcript length.
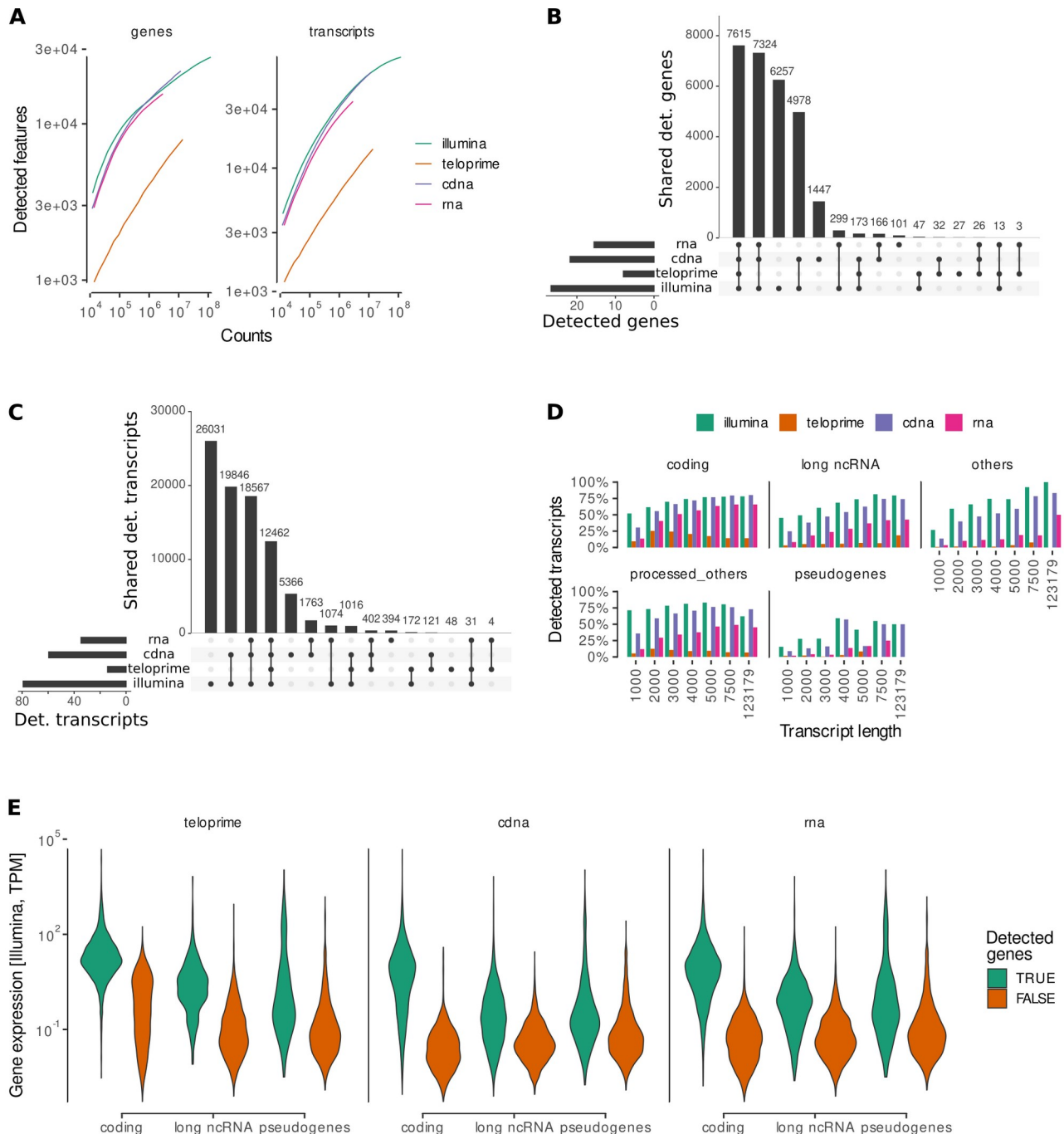

Fig. 3. Feature detection. **A** Feature detection rate by library size. A feature is counted as detected if there is at least one primary alignment to it. **B;C** Overlap of detected genes (B) and transcripts (C) between the different sequencing methods. **D** Transcript detection rate by transcript length and biotype. 100 % is any transcript detected in any of the sequencing datasets. E Abundance in the Illumina dataset of genes either detected or not by the different ONT library preparation methods.

Direct RNA- and cDNA-Seq is superior to TeloPrime for gene/transcript quantification

RNA-Seq and cDNA-Seq correlated very well on gene ($R^2 = 0.92$; Fig. 4) and transcript level ($R^2 = 0.92$; Fig. S4A). However, the situation was different when comparing the long- with short-read sequencing: While direct cDNA and RNA-Seq results correlated well with the abundance measured by Illumina sequencing on gene level ($R^2 = 0.85$ and 0.87), larger differences occurred on transcript level ($R^2 = 0.54$, both). The estimation of transcript abundance is challenging as transcripts from one gene share large parts of their sequence, causing ambiguity in read assignments when using short-reads (Soneson et al., 2016). TeloPrime-Seq quantification correlated less with the other methods. Noteworthy, the slope of the ratio of TeloPrime counts to those of other methods was larger than 1, indicating that the TeloPrime method overestimates the expression of highly abundant features and underestimates lowly expressed features, an observation also made in ONT RNA-Seq using PCR amplification (Chen et al., 2021). As we sequenced all TeloPrime and direct cDNA samples on two different flow cells, we could make use of technical replicates to assess the variability in sequencing performance between different flow cells. While sequencing counts for TeloPrime sequencing correlated very well among the two flow cells ($R^2 = 0.86$ to 0.90; Fig. S3B), the direct cDNA-Seq method showed a higher variation ($R^2 = 0.75$ to 0.77; Fig. S3C) reflecting the variability in read quality and length distributions (Fig. 1B-D). Thus, we find that direct RNA and direct cDNA protocols are most similar, reflecting a more unbiased representation of the transcriptome in comparison to TeloPrime.
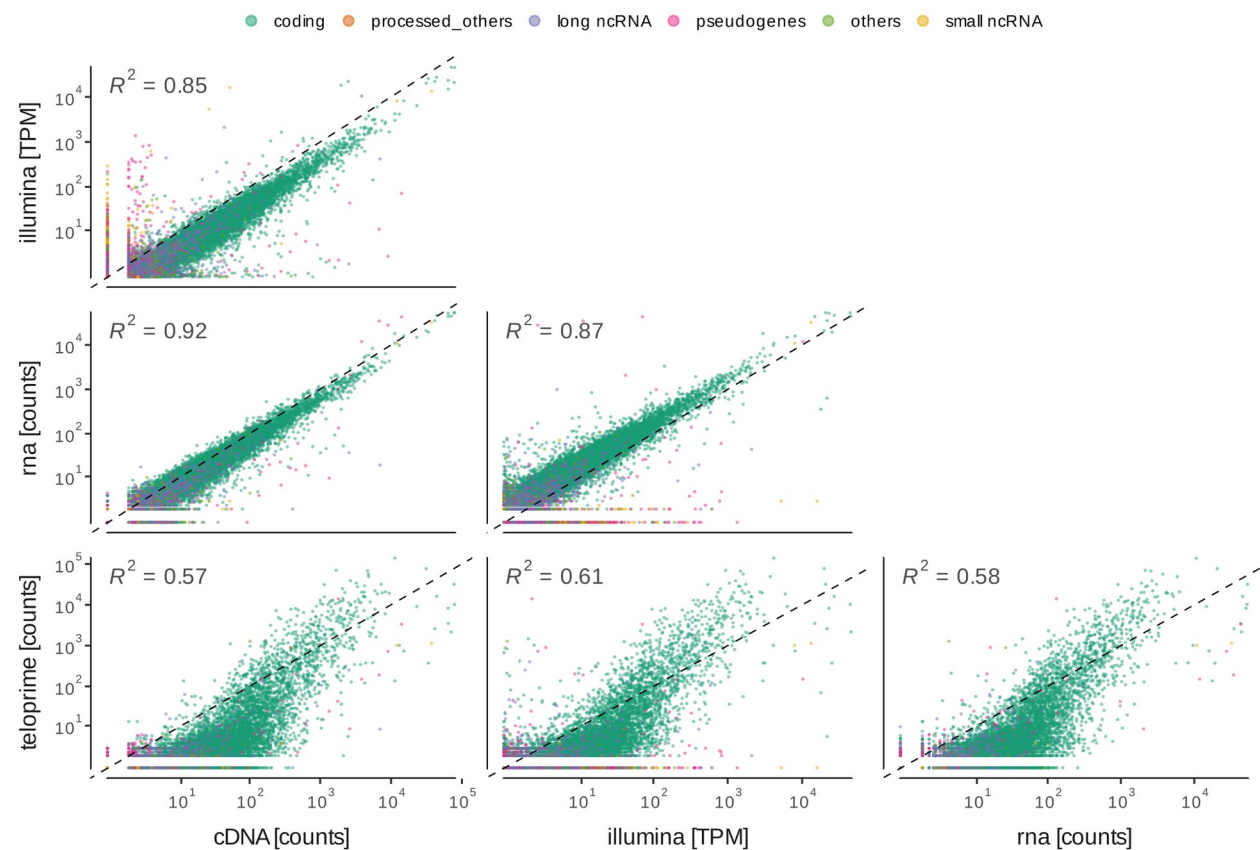
Fig. 4. Gene quantification. Scatter plots showing the correlation in gene quantification between the different sequencing methods.

Differential gene and transcript expression analysis in cold-activated BAT

The main goal of feature quantification is to detect differentially expressed genes and transcripts across biological samples. We compared the performance of the different ONT library preparation methods and Illumina sequencing to detect such features between iBAT of mice housed for 24 h either at room temperature or 4°C. Overall, the largest number of differentially expressed genes and transcripts (989 and 1195, respectively) was detected by Illumina sequencing followed by TeloPrime- (568 genes and 552 transcripts) and cDNA-Seq (489 genes and 476 transcripts) (Fig. 5A). Each sequencing method detected a unique set of features not seen by the other methods (312, 247 and 47 for Illimuna, TeloPrime and direct cDNA, respectively). Of note, irrespective of the method, most features identified as differentially expressed were protein coding genes. Interestingly, Illumina and direct cDNA performed better in detecting differentially regulated lncRNA genes compared with the TeloPrime protocol (Fig. 5B). Next, we compared the expression levels and fold changes between genes called to show differential gene or transcript expression by one of the ONT methods alone, Illumina sequencing alone or both methods. We found that genes differentially regulated according to Illumina but not the ONT

methods showed low expression in both types of analysis, but more evident in the long-read method (Fig. 5C). On the other hand, genes detected to be differentially regulated by the ONT methods showed similar expression levels in both short- and long-read sequencing, independent of their status according to Illumina-Seq, but those features called by both methods showed higher fold changes in Illumina-Seq upon cold treatment. Thus, we observed that TeloPrime-Seq showed higher fold changes compared to Illumina-Seq, confirming that this method overestimates highly expressed and underestimates lowly expressed features, and that transcripts (but not genes) called by either of the long but not short-read sequencing methods showed lower expression on average in Illumina-Seq (Fig. S4).
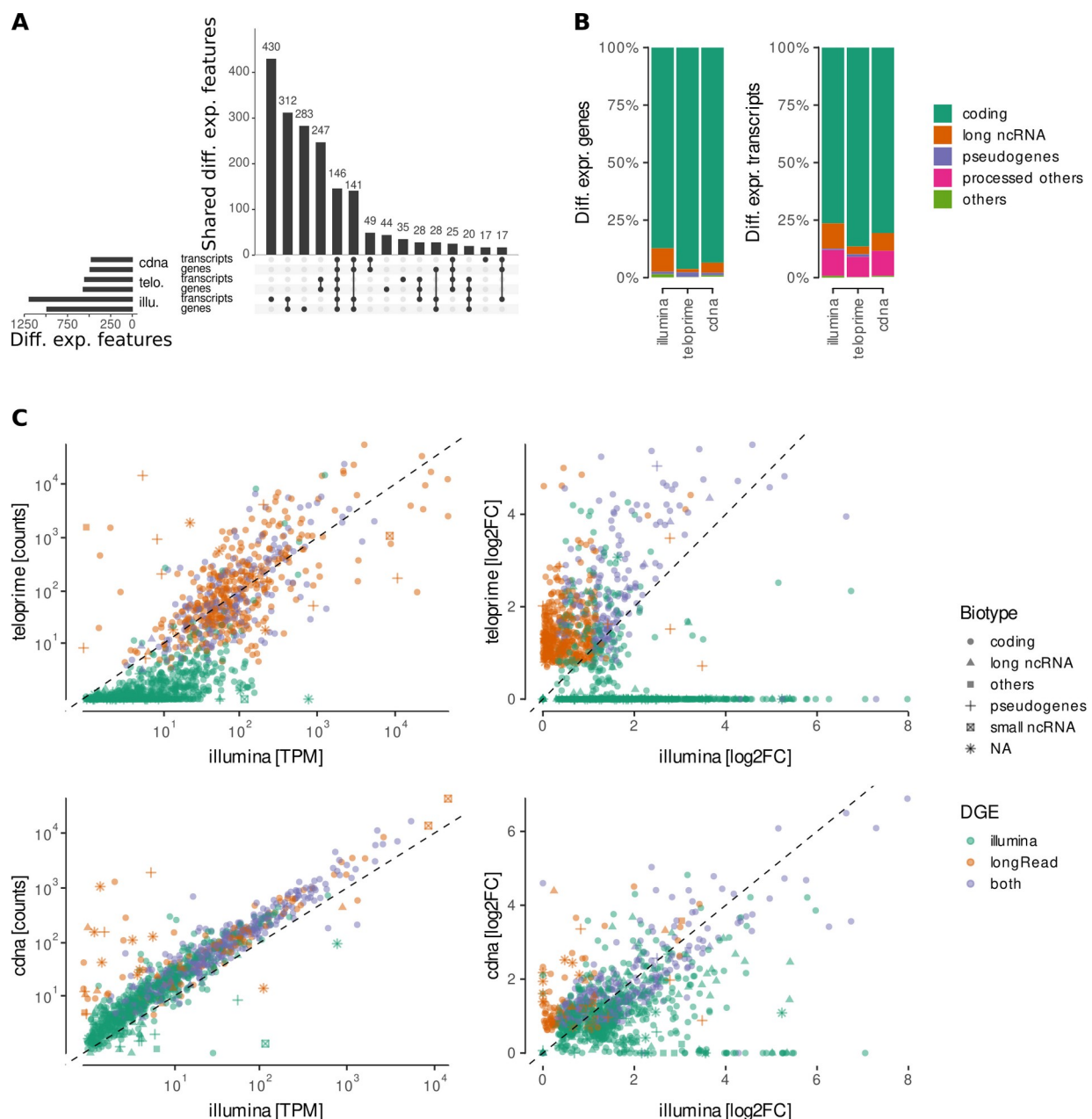
Fig. 5. Differential gene expression analysis. **A** Overlap between genes showing differential gene expression or genes with at least one transcript showing differential transcript expression over the two ONT methods and Illumina sequencing. **B** Biotypes of genes and transcripts showing differential gene and transcript expression respectively compared between the different sequencing methods. **C** Comparison of expression levels and fold changes of genes showing significant differential expression between direct cDNA/TeloPrime and Illumina sequencing.

ONT long read reannotation reveals novel features of the murine BAT transcriptome

The ability of long reads to unambiguously identify expressed isoforms facilitates the analysis of complex splicing events involving multiple exons. To reveal the nature and magnitude of newly identified transcripts in murine brown adipose tissue, we applied two transcriptome reassembly algorithms. FLAIR (Tang et al., 2020) corrects splice-sites of

long-reads based on known, user-provided annotation e.g., from short reads, filters for those long-reads starting at given transcription start sites (TSS) and then collapses the long-reads to transcripts, keeping those with a minimum coverage of ONT reads. Stringtie (Pertea et al., 2015; Shumate et al., 2021) creates a splice graph based on long-reads, moves the splice junctions in this graph to the nearest junctions supported by short-read sequencing, removing them if not supported, and uses both short and long reads to filter for a minimum coverage. Each identified transcript was assigned to a structural category describing the type of relationship to the reference transcript (Fig. 6A). Generally, stringtie reannotated more transcripts compared to FLAIR (Fig. 6A), as shown previously (Kovaka et al., 2019). Irrespective of the reannotation algorithm, direct cDNA sequencing yielded the highest, while direct RNA sequencing gave the lowest number of reannotated transcripts, resembling the number of mapped reads (Fig. 3B-D), suggesting that transcript identification is affected by the ONT sequencing protocol. In all stringtie and the direct RNA-Seq FLAIR reannotation, most reannotated transcripts fully matched reference annotations ("full splice match"; Fig. 6A). Novel transcripts not present in the reference annotation ("novel in catalog") made up for the second largest set and were relatively more prominent in FLAIR reannotations compared to stringtie reannotations, especially in the FLAIR-TeloPrime reannotation. Transcripts missing exons from either the 3' or 5' end (i.e. "incomplete splice match"; ISM) comprised the third largest class. Of note, the TeloPrime-(FLAIR) dataset was almost devoid of ISM transcripts, most likely because of its selective enrichment for full-length RNA molecules. Combinations of known splice junctions or splice sites were the prevailing mechanisms underlying transcript diversity among "novel transcripts" (Fig. 6B). Among the ISM, class 5' fragments as well as mono-exonsmatches", were most often found (Fig. 6C). Interestingly, ISMs were more prominent when the TeloPrime data were used for reannotation by stringtie, suggesting that the hybrid approach might reintroduce truncated isoform annotations potentially based on degraded RNA molecules. Noteworthy, even though only primary reads were used for the reannotation, the direct cDNA and the stringtie reannotation of the TeloPrime data, featured a substantial amount of antisense transcript annotations (Fig. 6A). These were on average shorter (Fig. S5A) and consisted of less exons compared to the full splice matching transcripts (Fig. S5B), indicating they might be artefacts. These annotations might interfere with transcript mapping, especially for non-directional sequencing methods. Comparison of the overlap of reannotated reference transcripts between the different datasets showed that the largest sets of transcripts were either detected by all combinations of reannotation algorithm and sequencing method, or only in all the stringtie datasets, highlighting the strong impact of the transcriptome reassembly method (Fig. 6D).

Intriguingly, large sets of transcripts were only detected in the teloprime/stringtie or in the Illumina dataset and to a smaller extent in the cDNA/stringtie and RNA/stringtie datasets. Among the FLAIR reannotations, only direct cDNA-Seq showed transcripts specific for this method. Curiously, while the TeloPrime/FLAIR dataset included the smallest amount of reference transcripts, the combination of TeloPrime and stringtie reannotated the highest number of reference transcripts apart from the Illumina sequencing-based reannotation. Thus, the ONT sequencing method had a significant impact not only on the number but also on the nature of the structural category of novel transcripts identified.
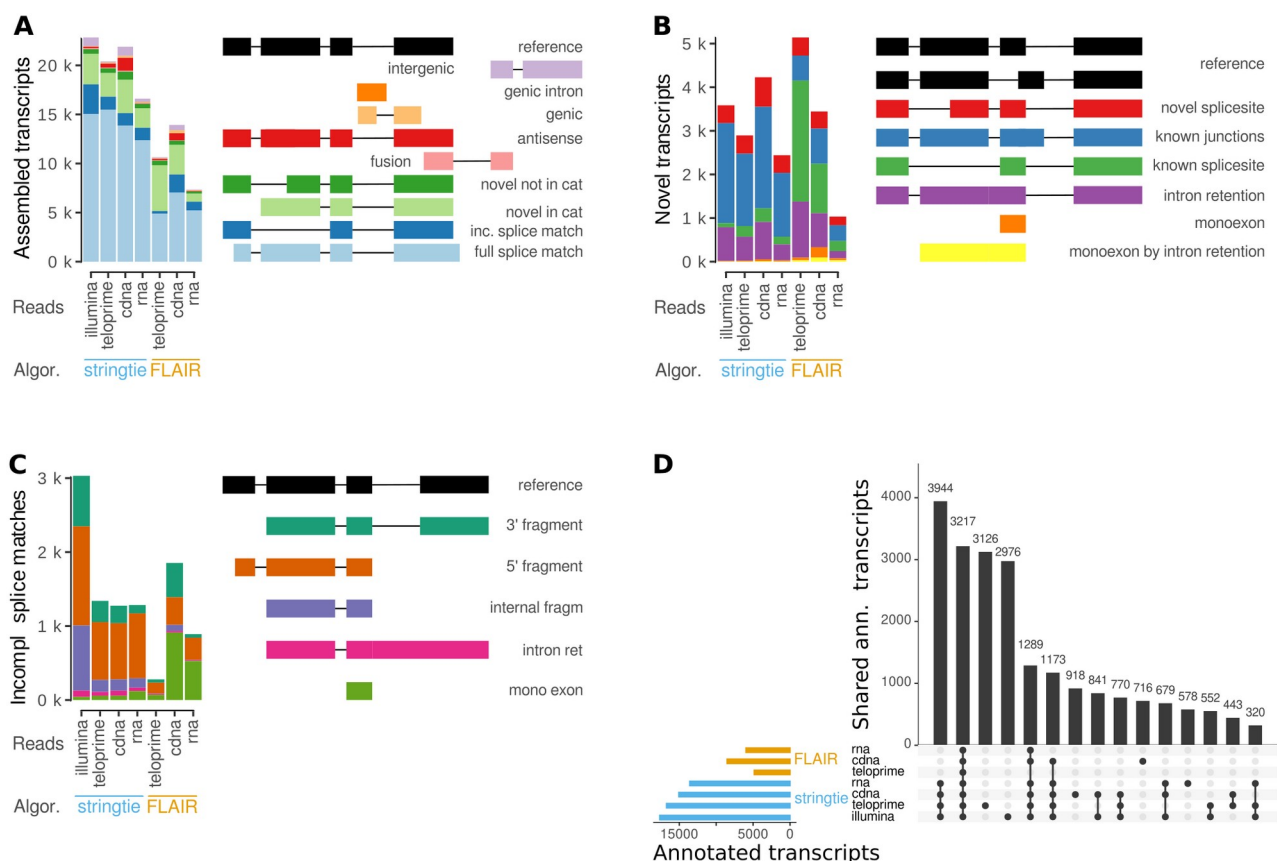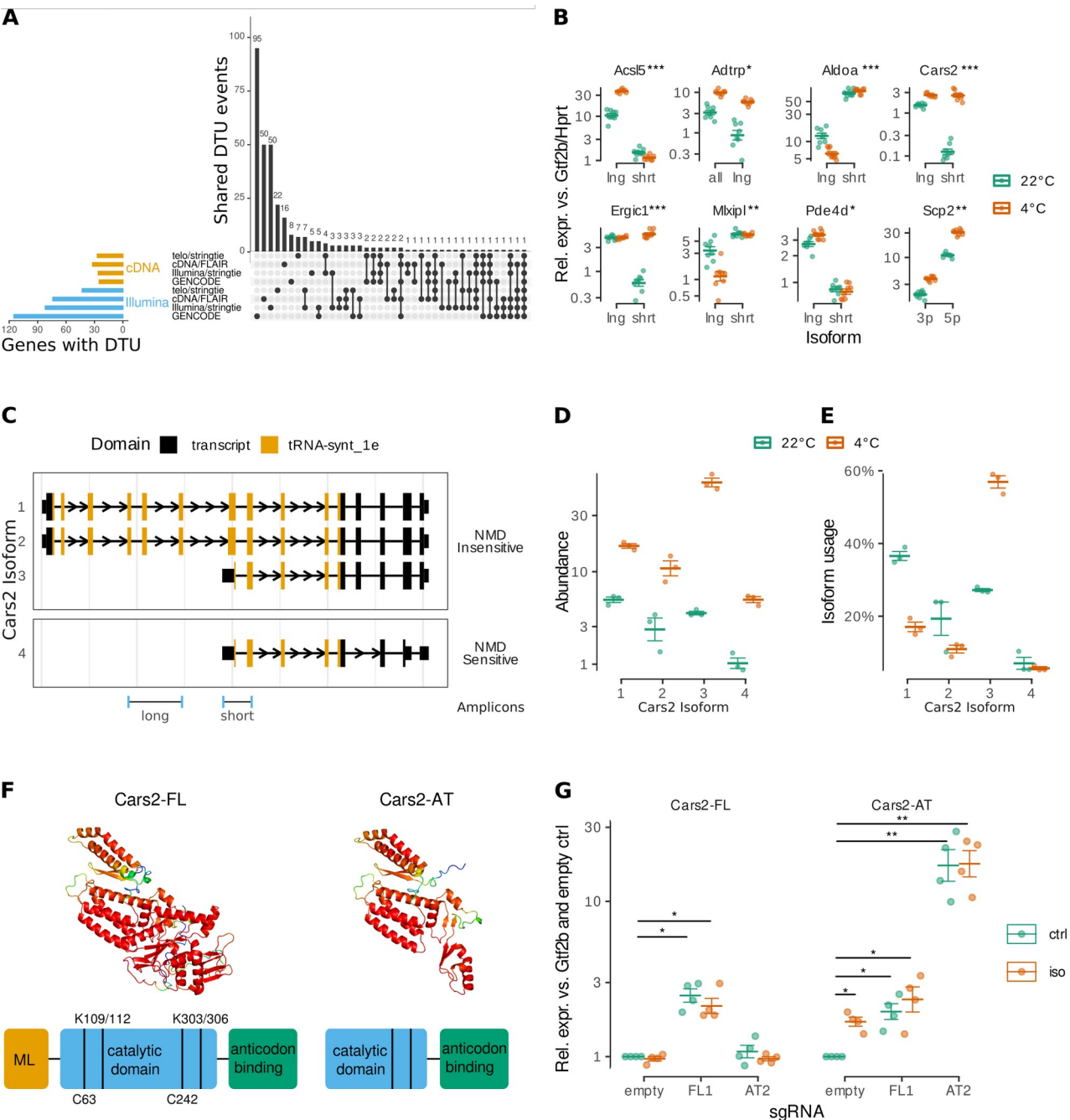

Fig. 6. Reannotation analysis of short and long -read sequencing protocols stratified by transcript assembler. **A** Overview of different structural categories. **B;C** Subclassification within the novel transcript and incomplete splice match categories. **D** Overlap between transcripts in the reference annotation (GENCODE M22) correctly reannotated by the different transcriptome assemblers and sequencing methods (includes transcripts falling into the full splice match and incomplete splice match categories as detailed in A).

Differential transcript usage analysis unravels gene expression alterations upon cold exposure in iBAT

Differential gene expression analysis lacks the sensitivity to detect changes at the transcript-level caused by e.g, alternative transcription start sites (TSS) or alternative splicing (De Paoli-Iseppi et al., 2021). To overcome this limitation, we applied differential transcript usage (DTU) analysis to identify genes using different transcripts in cold-

activated compared to inactive iBAT. Reliable identification of DTU depends critically on both the accuracy of the transcript expression quantifications as well as the transcriptome annotation. Therefore, we investigated combinations of transcript quantitation (Illumina or
355 direct cDNA counts) and reannotation algorithms (Fig. 7A). As observed for the differential gene expression analysis (Fig. 5A), a higher number of DTUs were identified when Illumina counts were used to assess transcript quantification (Fig. 7A). We found little overlap between stringtie and FLAIR reannoations in line with other reports (Ringeling et al., 2022). Among the genes with significant isoform switches between cold-activated BAT
360 compared to the controls, we observed phosphodiesterase 4D (*Pde4*d), regulating levels of the signalling intermediate cAMP, which activates lipolysis, glucose uptake, and thermogenesis in brown adipocytes (Reverte-Salisa et al., 2019); the thermogenesis regulating hydrolase androgen dependent TFPI regulating protein (*Adtrp; P. Li et al., 2022*) and regulators of fatty acid metabolism (acyl-CoA synthetase long-chain family member 5;
365 *Acsl5*), glycolysis (Aldolase A; *Aldoa*), protein sorting (endoplasmic reticulum-golgi intermediate compartment 1; *Ergic*1), lipid synthesis (MLX interacting protein-like; *Mlxip*l), beta-oxidation (sterol carrier protein 2, liver; *Scp2*) and protein cysteinylation (cysteinyl-tRNA synthetase gene; *Cars2; Akaike et al., 2017;* Fig. 7C-E and S7). qPCR analysis on iBAT from control and cold-treated mice using primer sets specific for the individual
370 transcripts corroborated our DTU analysis, thus validating the isoform regulation (Fig. 7B). Incorporation of histone modification chromatin immunoprecipitation (ChIP)-Seq peaks into the FLAIR algorithm revealed that *Cars2* is transcribed from two different TSS, i.e. the 'canonical' TSS and an alternative gene-internal TSS, giving rise to either the full-length transcript (*Cars2*-FL) or a 5' truncated transcript (*Cars2*-AT; Fig. S6). Visual inspection of
375 ONT direct cDNA sequencing coverage data indicated that the canonical TSS is dominant at room temperature, while the the alternative TSS is predominantly used in iBAT of cold treated animals (Figure S6A). Cars2 is the mitochondrial cysteinyl-tRNA synthetase, which is important for the translation of mitochondrially encoded genes (Rajendran et al., 2018), but additionally executes a 'non-canonical' function in post-translational cysteine and
380 protein persulfidation ultimately affecting mitochondrial respiration (Akaike et al., 2017). Therefore, we analysed the sequences of *Cars2*-FL and *Cars2*-AT to identify whether these alternative transcripts have coding potential using the *Coding-Potential Assessment* (CPA) Tool. CPAT predicted both isoforms to be coding (coding potential > 0.99), annotating the open reading frame correlating with the UniProt reference amino acid
385 sequence to Cars2-FL and predicting a N terminally truncated protein isoform for Cars2-AT using the same open reading frame but missing the first 244 aa (Fig. 7F). As expected for a mitochondrial protein, targetP predicted a mitochondrial localisation signal at the N

terminus of the full length CARS2 protein, which was missing in the truncated protein isoform (Fig. 7F). The truncated CARS2 protein isoform was further predicted to lack parts

390 of the conserved binding sites for both Zn^{2+} and pyridoxal phosphate (PLP), indicating a potential lack of catalytic function (Fig. 7F). Using single guide RNAs (sgRNA) targeting either the canonical or the alternative promoter we aimed to specifically overexpress Cars2-FL and Cars2-AT in wt1-SAM immortalised brown adipocytes (Fig. 7G). Using a sgRNA targeting the alternative promoter, it was possible to induce the expression of

395 *Cars2*-AT 10 to 30 times, both with and without β-adrenergic stimulation (p = 0.005 and 0.003) without confounding effects on *Cars2*-FL expression (Fig. 7G). Thus, these results demonstrate that long-read DTU can identify and quantify biologically relevant changes in isoform usage.

Fig. 7. A Overlap DTU events detected by different combinations of datasets used for transcript annotation and sequencing. **B** qPCR validation of selected DTU events in murine iBAT from control (22°C) or cold (4°C) treated mice. **C** Structure of expressed *Cars2* isoforms. Arrows show direction of transcription. Narrow lines show intronic regions (not to scale). Exons displayed as boxes. Taller exonic boxes are coding regions, shorter boxes are 5′ and 3′ UTR regions. Colours represent identified protein domains. **D;E** Abundance (normalised counts) and isoform usage of expressed *Cars2* isoforms. **F** Schematic representation of the general structure of Cars2 showing domains and residues important for catalytic activity (modified from (Akaike et al., 2017) and prediction of secondary and tertiary structures of the full length CARS2 and the predicted truncated protein isoform by LocalColabFold. The colour code depicts model confidence; green is low, red is high. **G** Expression of Cars2 isoforms in wt1-SAM brown adipocytes after targeting transcript specifics promoters using sgRNAs.

6.   Discussion

In this study, we assessed three different ONT long-read sequencing protocols as well as Illumina short-read sequencing for differential gene expression, transcriptome reannotation and differential transcript usage analysis giving us unprecedented views on transcript diversity in murine brown adipose tissue. While studies of differentially expressed genes have provided much of our current understanding of molecular mechanisms controlling BAT function (Seale, 2015; Shapira & Seale, 2019; W. Wang & Seale, 2016), determining genes where functionally distinct alternative transcripts change between BAT activity states is an interesting gene-regulatory mechanism to uncover. Feature quantification using direct RNA and direct cDNA protocols correlated well with Illumina on gene level, and to a lesser extent on transcript level, as described (Sessegolo et al., 2019). Direct RNA and direct cDNA sequencing showed an even higher correlation on gene level which was only slightly reduced when quantifying transcripts, suggesting that *(i)* reverse transcription to cDNA has a limited impact on transcript quantification and *(ii)* long reads give better estimates of transcript abundances, as they more often unambiguously map to a single transcript. In fact, direct RNA as well as direct cDNA even outperform Illumina in terms of accuracy of transcript quantification and differential expression, which has been attributed at least in part to the lack of GC content bias (Chen et al., 2021; Oikonomopoulos et al., 2016, 2020; Sessegolo et al., 2019). The key challenges with direct RNA sequencing are the large amount of input RNA required, higher error rate as compared with cDNA sequencing and the lack of multiplexing options (Chen et al., 2021; Soneson et al., 2019). PCR amplification protocols such as the TeloPrime typically produce higher sequencing depth than PCR-free methods, increasing coverage which is required for accurate identification of alternative transcripts (Chen et al., 2021; Glinos et al., 2022). However, the TeloPrime method overestimated the abundance of highly expressed and underestimated the abundance of lowly expressed features, caused in part by the inherent PCR amplification step compromising transcript diversity (Chen et al., 2021; Sessegolo et al., 2019).

Comparison of the transcriptome reannotation methods showed that when using the same input data, stringtie surpasses FLAIR in terms of the number of correctly reassembled transcripts from the reference annotation (full and incomplete splice matches; Fig. 6A), in line with previous reports (Kovaka et al., 2019). In contrast to stringtie, FLAIR incorporates ChIP/CAGE-Seq data which can be high value because it discriminates between true internal TSS, and artefacts from 5' degraded RNA, which importantly allowed us to identify a novel 5' truncated *Cars2* transcript isoform (*Cars2*-AT) highly induced in iBAT of cold-

435   treated mice and predicted to encode a N-terminally truncated protein (Fig. S6). Cars2 has recently been reported to be involved in sulphur metabolism, which is of importance for mitochondrial morphology and BAT function (Akaike et al., 2017; Soriano et al., 2018). The functional significance of the novel *Cars2*-AT reported here, the predicted changes in localization, and its role in thermogenesis remain to be experimentally defined. Our brown

440   adipocyte cell model overexpressing *Cars2*-AT and *Cars2*-FL by CRISPR/Cas9 mediated activation of the respective endogenous promoter (Fig. 7G) will be a valuable tool to answer these questions; and will also allow us to test whether Cars2-AT may have a dominant negative regulatory role in Cars2 expression, as observed for other truncated protein isoforms (Gervois et al., 1999; Tomita et al., 2003). TeloPrime's strategy for

445   enrichment of full-length transcripts allowed us to identify several novel alternative transcripts produced from the same gene with presumably important functional consequences on protein structure, culminating on thermogenic β3-adrenergic receptor (AR) mediated cAMP signaling (Fig. 7B and C): Cellular cAMP levels are also regulated cAMP-specific phosphodiesterases (PDEs) and *Pde4d* regulates lipolysis and thermogenic

450   gene expression (Vezzosi & Bertherat, 2011). Cold-activated induction of *Pde4d*-long might lead to reduced thermogenic cAMP signalling since *Pde4d*-long controls cAMP levels negatively in a spatial manner due to UCR motifs present in the long *Pde4d* isoform (Lynch et al., 2007). Another factor regulating the thermogenic β3-AR signalling cascade is *Adtrp*; *Adtrp*-deficient mice are cold-intolerant and have defective thermogenesis (P. Li et

455   al., 2022). Thus, cold-activated induction specifically of the longer (enzymatically active) Adtrp isoform shown here might enhance BAT function. Peroxisome derived lipids are required for brown fat-mediated thermogenesis through regulation of cold-induced mitochondrial fission (Kleiboeker & Lodhi, 2022). The cold-induced switch in the peroxisome lipid transfer protein *Scp2* in favour of the longer (and enzymatically active)

460   isoform (Fig. 7B) may thus support a peroxisome-to-mitochondria lipid signalling hub, supporting mitochondrial uncoupling and thermogenesis (N. C. Li et al., 2016). Proper BAT thermogenic function requires cellular protein quality control and removal of misfolded proteins (Bartelt et al., 2018). We here identified DTU in the protein sorting gene *Ergic1 (Joshi et al., 2017)* with the shorter isoform being specifically increased upon cold activation

465   (Fig. 7B;C) which may thus participate in the homeostatic adaptation of BAT to cold stress involving the ER stress response.

The key transcription factor regulating *de novo* lipogenesis *Mlxipl (a.k.a. ChREBP*) inhibits BAT thermogenesis and downregulates expression of genes involved in mitochondrial biogenesis and respiration (Wei et al., 2020). Based on our DTU analysis of *Mlixpl*, we

470 speculate that the observed downregulation of the long, intact *Mlixpl* isoform specifically in cold-activated iBAT alleviates *Mlixpl* inhibitory effects of on BAT thermogenesis. Fatty acid oxidatiocyl-CoA synthetases such as *Acsl5* regulate fatty acid trafficking and metabolism (Bowman et al., 2016). Acsl5 activity increases adiposity, decreases Ucp1 expression and energy expenditure in mice (Bowman et al., 2016). Here, we show DTU in *Acsl5* giving rise

475 to two transcripts differing in their 5'UTR but otherwise identical protein domain structure (Fig. 7B;C) with the longer isoform being the almost exclusively expressed variant in cold, suggesting differential control of translation efficiency of this fatty acid channelling enzyme in response to cold challenge in BAT. A similar scenario may be true for the glyceroneogenic enzyme Aldoa, which controls the cellular levels of glycerol-3-phosphate

480 (G3P) shown to be increased upon cold exposure in BAT of mice (Moura et al., 2005). Thus, a systematic characterization of isoform-level variation and complexity in activated BAT as described here will help understand how isoforms might contribute to the regulation of BAT function.

7. Material and methods

Animal experiments
485

Unless otherwise stated, mice were kept at 22°C to 24°C on a regular 12 h light cycle with *ad libitum* access to food (Altromin 1324, Altromin Spezialfutter GmbH & Co. KG, Lage, Germany) and water. 20 weeks old male wild type C57BL/6N mice were singly housed at 4 °C for a period of 24 h prior to harvesting adipose tissues.

RNA isolation
490

Whole frozen iBAT samples were homogenised in 1 ml TRIsure (Bioline, Memphis, Tennessee, USA) per animal using a tabletop homogeniser (FastPrep-24 5G, MP Biomedicals, Irvine, California, USA). RNA was isolated by phenol chloroform extraction and alcohol precipitation as described by (Workman et al., 2019).

Illumina RNA sequencing
495

The NEBNext Ultra II Directional RNA Library Prep Kit (New England BioLabs, Ipswich, Massachusetts, USA) was used to prepare 50 nt paired-end, strand specific libraries following the manufacturer's protocol and sequenced on a NovaSeq 6000 (Illumina Inc., San Diego, California, USA) for approximately 25 million reads per library.

ONT library preparation

For all experiments, sequencing on the GridION platform (ONT, Oxford UK) was performed using FLO-MIN106 R9 flowcells (ONT). Libraries prepared according to the TeloPrime and the direct cDNA protocol were sequenced on two different FLO-MIN106 R9 flow cells to examine sequencing variability.

poly(A) enrichment

RNA used for ONT sequencing was poly(A+) selected in two consecutive rounds using oligo(dT) beads (GenElute mRNA Miniprep Kit, Sigma MRN10, MilliporeSigma, Burlington, Massachusetts, USA) following the manufacturer's recommendations. Subsequently, RNA was alcohol precipitated using sodium acetate and glycogen following the protocol from the Ribo-Zero rRNA Removal Kit (Illumina).

TeloPrime libraries

The TeloPrime Full-Length cDNA Amplification Kit (Lexogen, Vienna, Austria) was used to select for full length mRNAs with intact 5' CAPs from 7 ng poly(A+) RNA. The resulting cDNA was PCR amplified with SYBR Green I (MilliporeSigma), TeloPCR enzyme mix and 3' and 5' primers (RP: 5'-TCTCAGGCGTTTTTTTTTTTTTTTTTT-3' and FP: 5'-TGGATTGATATGTAATACGACTCACTATAG-3') to determine the optimum cycle numbers for the large-scale PCR to generate enough material for long-read sequencing. The determined cycle number of 27 was applied for large scale PCR in the absence of SYBR Green I followed by processing of 400 ng of the cDNA with the SQK-LSK109 ligation sequencing kit (ONT) and the EXP-NBD104 barcoding kit (ONT) following manufacturer's instructions.

Direct cDNA libraries

Libraries were prepared from 100 ng poly(A+) RNA using the SQK-DCS109 direct cDNA sequencing kit (ONT) and the EXP-NBD104 barcoding kit (ONT) according to manufacturer's protocol.

Direct RNA libraries

Libraries were prepared from 500 ng poly(A+) RNA using the SQK-RNA002 direct RNA sequencing kit according to manufacturer's protocol (ONT).

Reverse transcription and qPCR

530 RNA was reverse transcribed into cDNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems 4368814, Applied Biosystems, Waltham, Massachusetts, USA) following the manufacturer's instructions. qPCR was performed in 384 well format in a LightCycler 480 II (Roche, Basel, Switzerland). 4 µl of 1:20 diluted cDNA, 0.5 µl gene specific primer mix (5 µl each) and 4.5 µl FastStart Essential cDNA

535 Green Master (Roche) were amplified using 45 cycles of 25 s at 95 °C, 20 s at 58 °C and 20 s at 72 °C after 300 s at 95 °C initial denaturation. All combinations of primers and samples were run in duplicates and C_q values calculated as the second derivative maximum. Genes of interest were normalised against housekeeper genes using the ΔC_q method. The primers used in this study are listed in Tab. S1.

Long-read alignment and quantification

540

For the reference-based comparison of ONT library preparation methods, reads were mapped against the transcriptome (GENCODE M22; `-ax map-ont --secondary=no -uf`) and genome (GRCm38.p6; `-ax splice --secondary=no -uf`) using minimap2 (H. Li, 2018). Reads were subsequently filtered for an average PHRED score > 7 using

545 NanoFilt (De Coster et al., 2018). GenomicAlignments (Lawrence et al., 2013) was used to directly extract read level information from BAM files, including transcript level reference sequence, flag based mapping types, mapping position and cigar based aligned length information as described by (Soneson et al., 2019).

Short-read alignment and quantification

550 Illumina short-reads were quality filtered using cutadapt (Martin, 2011; `-q 28 -m 30`) and mapped to the genome (GRCm38.p6) using STAR (Dobin et al., 2013). Transcript level quantification was done using salmon in selective alignment mode (GENCODE M22; Patro et al., 2017). Gene level abundance estimation was done using tximport (Soneson et al., 2016).

Differential gene and transcript expression analysis

555

Differential gene and transcript expression analysis was done using DESeq2 with H_0: log2FC > 0.5 (Love et al., 2014). Adjustment for multiple testing was done using the *s*-value method proposed by Stephens (2017) implemented in apegeglm (Zhu et al., 2018) and *s*-values < 0.05 were considered significant.

Transcriptome reannotation

Transcriptome reannotation using stringtie (Shumate et al., 2021) was done with a splice junction cutoff (-j) of 10 and in presence of the reference annotation (-G, GENCODE M22). For ONT long-read runs, the --mix mode was used, additionally providing stringtie with the Illumina-Seq data of the same sample. Afterwards, all reannotations were pooled within the respective library preparation methods using stringtie --merge with a coverage cutoff (-c) of 3 and only isoforms with a minimum isoform fraction of 5 % per gene were kept (-f). No reference annotation was used in the merge step.

Aligned ONT long-reads were corrected by FLAIR correct (Tang et al., 2020) using the splice junction data from STAR from the Illumina-Seq runs of the same samples, filtered for a minimum splice junction coverage of 10. Subsequently, the corrected reads were collapsed into a transcriptome reannotation by FLAIR collapse using a minimum coverage of 3 and setting the --stringent flag. In order to mark true transcriptional start sites, a combined bed file from FANTOM5 CAGE peaks and iBAT H3K4me3 peaks was provided (Abugessaisa et al., 2017; Engelhard et al., 2022). For the TeloPrime data, the --trust_ends flag was additionally set.

The reannotated transcripts were compared to the reference annotation using SQANTI2 (Tardaguila et al., 2018)and the overlap between the different sequencing methods was calculated based on the associated transcripts for transcripts either fully or partially matching a reference transcript as described elsewhere in detail (Soneson et al., 2019).

Analysis of differential transcript usage

Only genes with at least 10 counts in all samples were included in the analysis. Transcripts were filtered for a minimum of 5 counts and 10 % of the counts of the parent gene in half of the samples. Differential transcript usage analysis was done using DRIMSeq and FDR calculated using stageR (Love et al., 2018; Nowicka & Robinson, 2016; Van den Berge et al., 2017). FDR < 0.1 and transcript usage changes > 5 % were considered significant. In order to detect differential transcript usage from qPCR data, a linear model was fitted with isoform and temperature as variables and the interaction term was tested. *p*-values were adjusted for multiple testing using Holm's method.

Annotation of transcripts

590   Annotation of transcript isoforms including open reading frames, nonsense mediated decay (spliceR, Vitting-Seerup et al., 2014) and functional protein domains (pfam, Finn et al., 2016) was done using IsoformSwitchAnalyzeR (Vitting-Seerup & Sandelin, 2019).

Cars2 protein fold predictions were generated in LocalColabFold using standard parameters (Mirdita et al., 2022). Computation of the models was performed on the
595   UCloud interactive HPC system, which is managed by the eScience Center at the University of Southern Denmark.

Cars2 overexpression in cell culture

wt1-SAM brown preadipocytes were grown in high glucose DMEM supplemented with 10 % fetal bovin serum (FBS) and 1 % penicillin-streptomycin. After reaching confluency,
600   differentiation was induced by 0.5 µM rosiglitazone, 1 nM T3, 1 µM Dexamethasone, 850 nM insulin, 125 µM indomethacine and 500 µM IBMX. Two days later, medium was exchanged for medium supplemented with 0.5 µM rosiglitazone and 850 nM insulin. Afterwards, medium was changed for medium containing 0.5 µM rosiglitazone every second day until reaching full differentiation 7 days after induction.

605   For *in vitro* gain of function studies using the wt1-SAM cell line, single guide RNAs (sgRNAs) were designed using CRISPick (Doench et al., 2014) and cloned into the sgRNA(MS2) cloning backbone (addgene 61424) as described by (Konermann et al., 2015). To transfect mature adipocytes, 3 µl TransIT and 250 ng plasmid DNA or 1.4 µl (10 µM) in 100 µl Opti-MEM I were pipetted into a well of a 24 well plate. After 15 min,
610   500,000 cells resuspended in 500 µl Opti-MEM I were added. 24 h later, medium was changed for regular differentiation medium. Guide sequences used were ATTTAGGCATTTGGGCACGG for Cars2-FL and GTGGCTGAACAGATCTGGCC for Cars2-AT.

8.   Data and software availability

The RNA-Seq data presented in this study has been deposited to GEO and will be made
615   public once this study has been accepted for publication. Reviewers can get access via a reviewer token upon request. Source code for the reproduction of all analyses and plots in this study are available at https://gitlab.com/kiefer.ch/nanoporeibat_new.

9.   Acknowledgements

10.   Conflict of interest

The authors declare no competing interest.

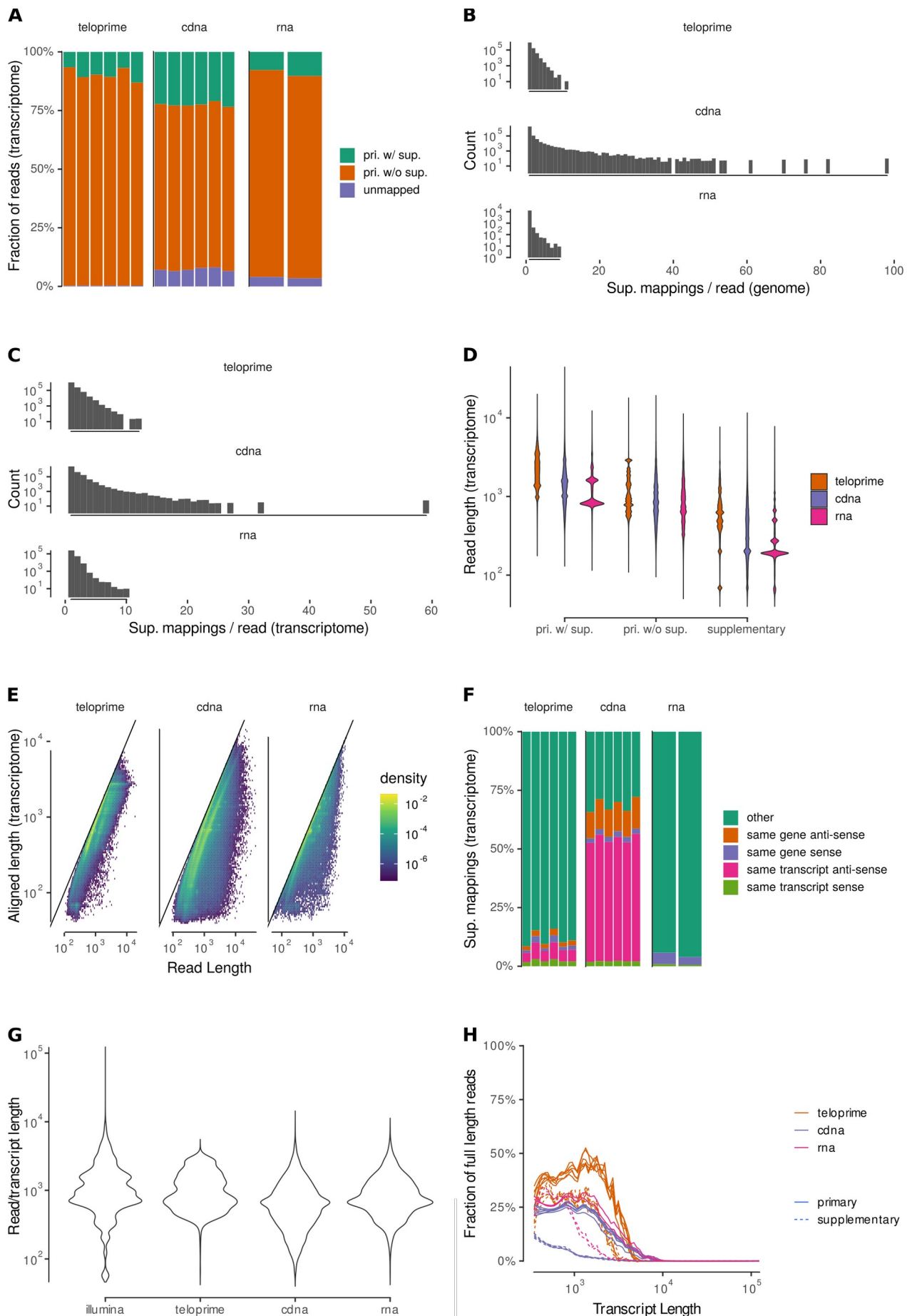630 **11.   Supplementary Figures**

Fig S1: Characterisation of read alignments. **A** Fraction of reads classified by whether the primary alignment against the transcriptome has at least one supplementary alignment attached to it. **B;C** Distribution of supplementary alignments per primary read for genome (B) and transcriptome (C) mapping. **D** Read length distribution for reads aligned to the transcriptome. **E** Aligned length vs read length for primary alignments to the transcriptome. **F** Fraction of supplementary alignments against the transcriptome stratified by their relation to the corresponding primary alignment. **G** Read length distribution of primary alignments to the transcriptome compared to the hypothetical transcript length distribution as estimated from the short-read transcript abundances. **H** Smoothed fraction of full length alignments (> 90 %) mapped to the transcriptome for transcripts > 350 nt.
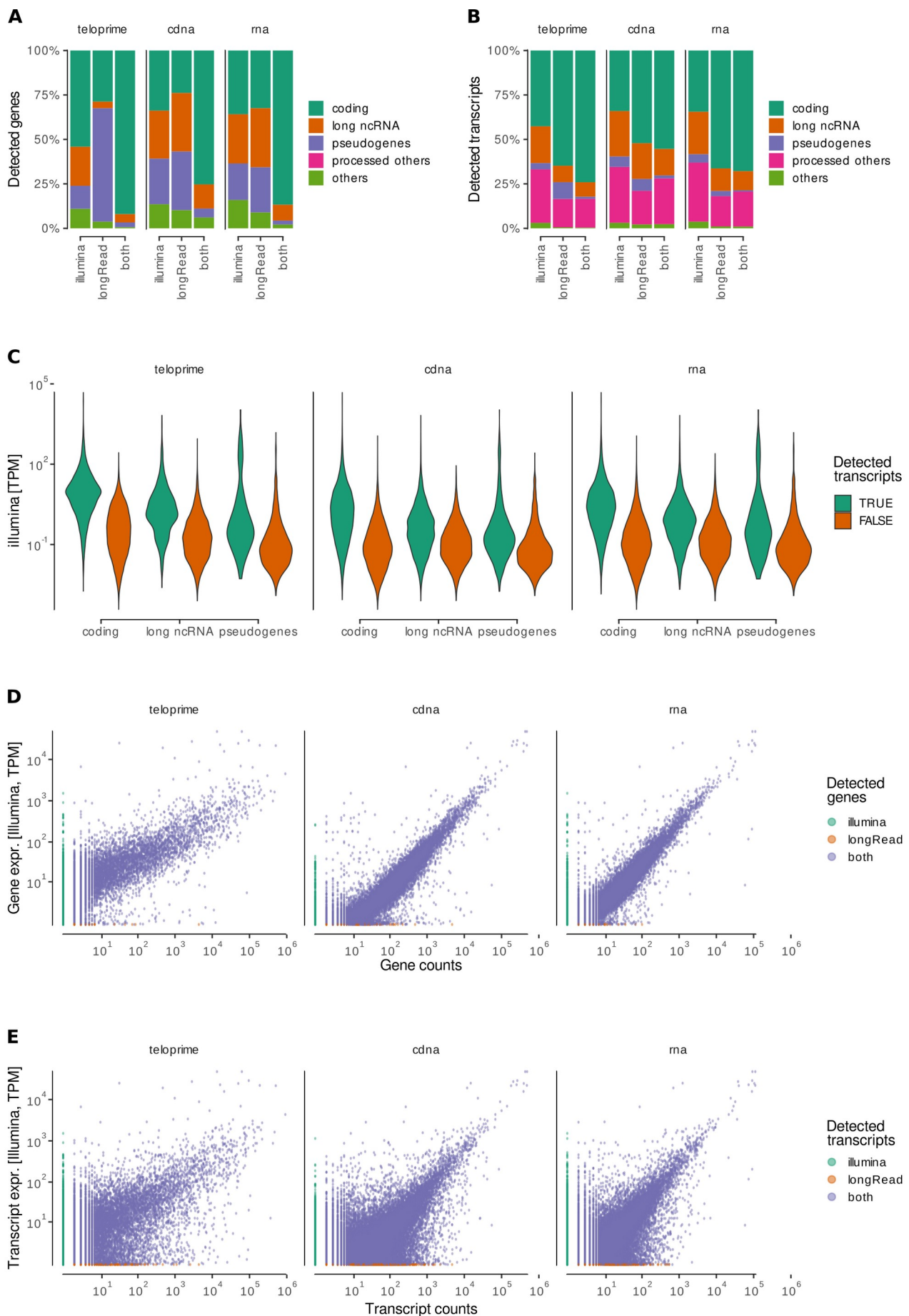
Fig S2: Feature detection. **A;B** Biotypes of genes (A) and transcripts (B) detected by the indicated method. A feature is counted as detected if there is at least one primary alignment to it. **C** Illumina-Seq based

645 abundance of transcripts stratified by whether they are detected by the different ONT library preparation methods or not. **D;E** Expression levels of genes (D) and transcripts (E) stratified by whether they are detected by Illumina, the respective ONT library preparation method, or both.
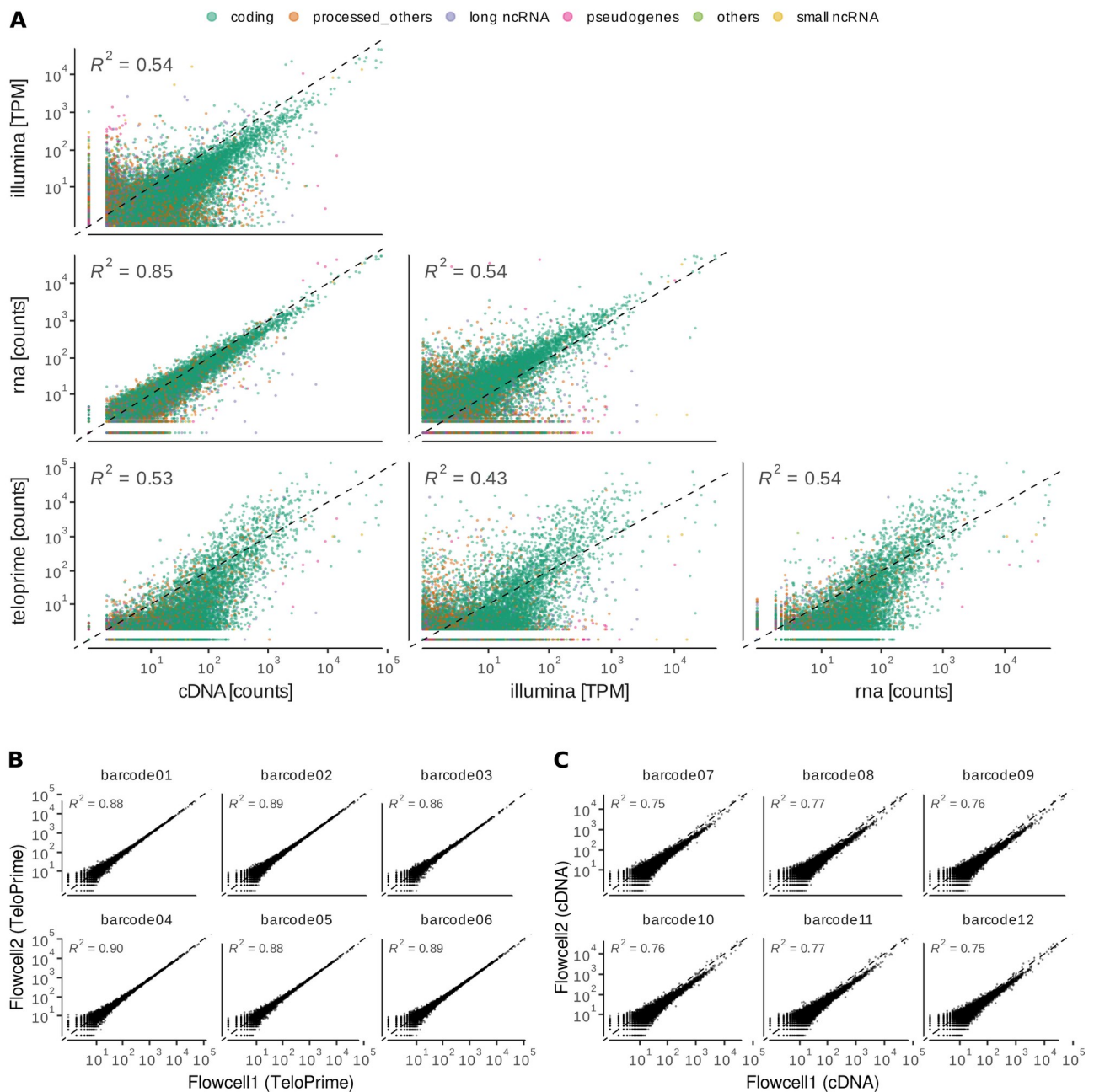
Fig S3: Transcript quantification. **A** Scatter plots showing the correlation in transcript quantification between the different sequencing methods. **B;C** Sample wise correlation analysis of transcript quantification between two different flow cells for the TeloPrime (B) and direct cDNA (C) methods.
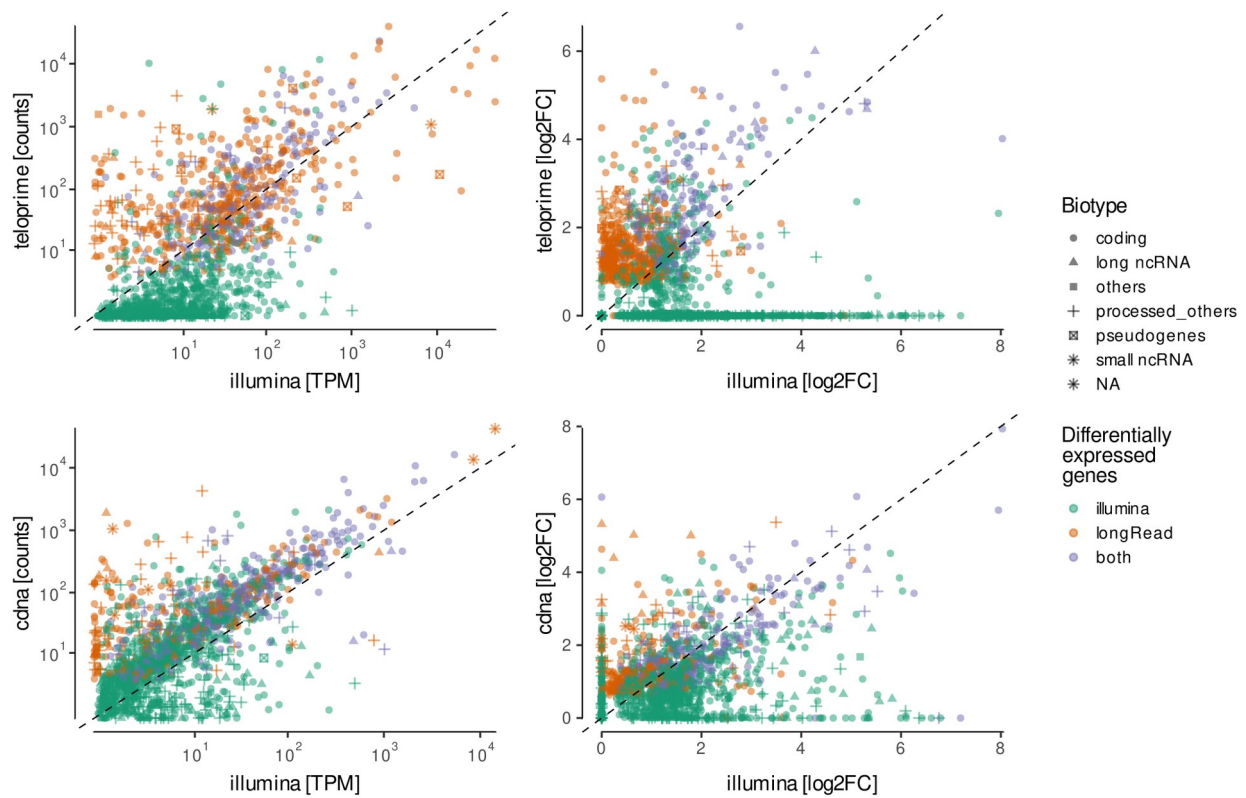
650

Fig S4: Differential transcript expression analysis. Comparison of expression levels and fold changes of transcripts showing significant differential expression between direct cDNA/TeloPrime and Illumina sequencing.
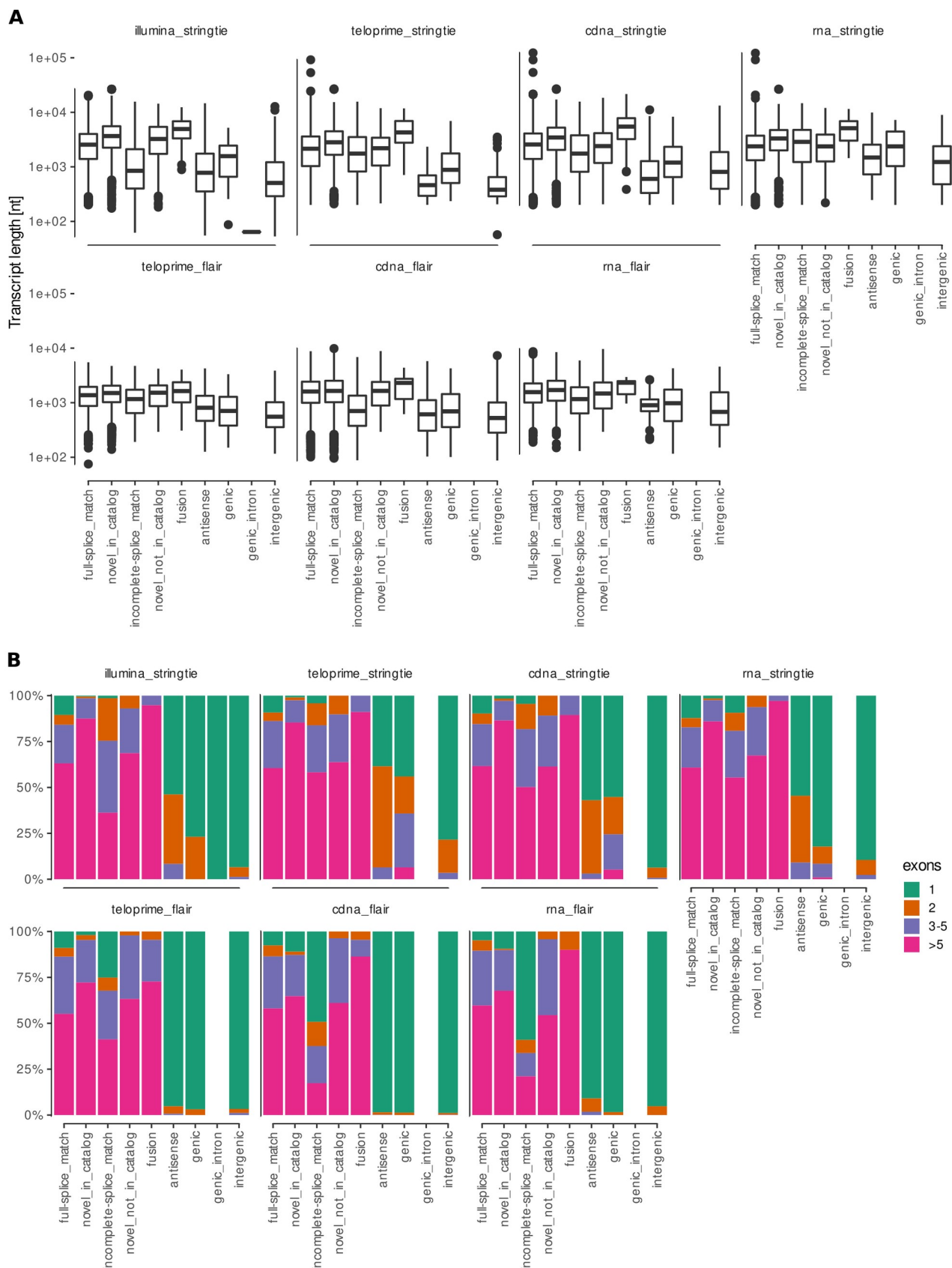
Fig S5: Reannotation Analysis. Transcript length distributions (A) and distributions of number of exons per reannotated transcript (B) compared for different structural categories and different reannotation methods. Center lines represent the median; hinges represent first and third quartiles; whiskers the most extreme values within 1.5 interquartile range from the box.
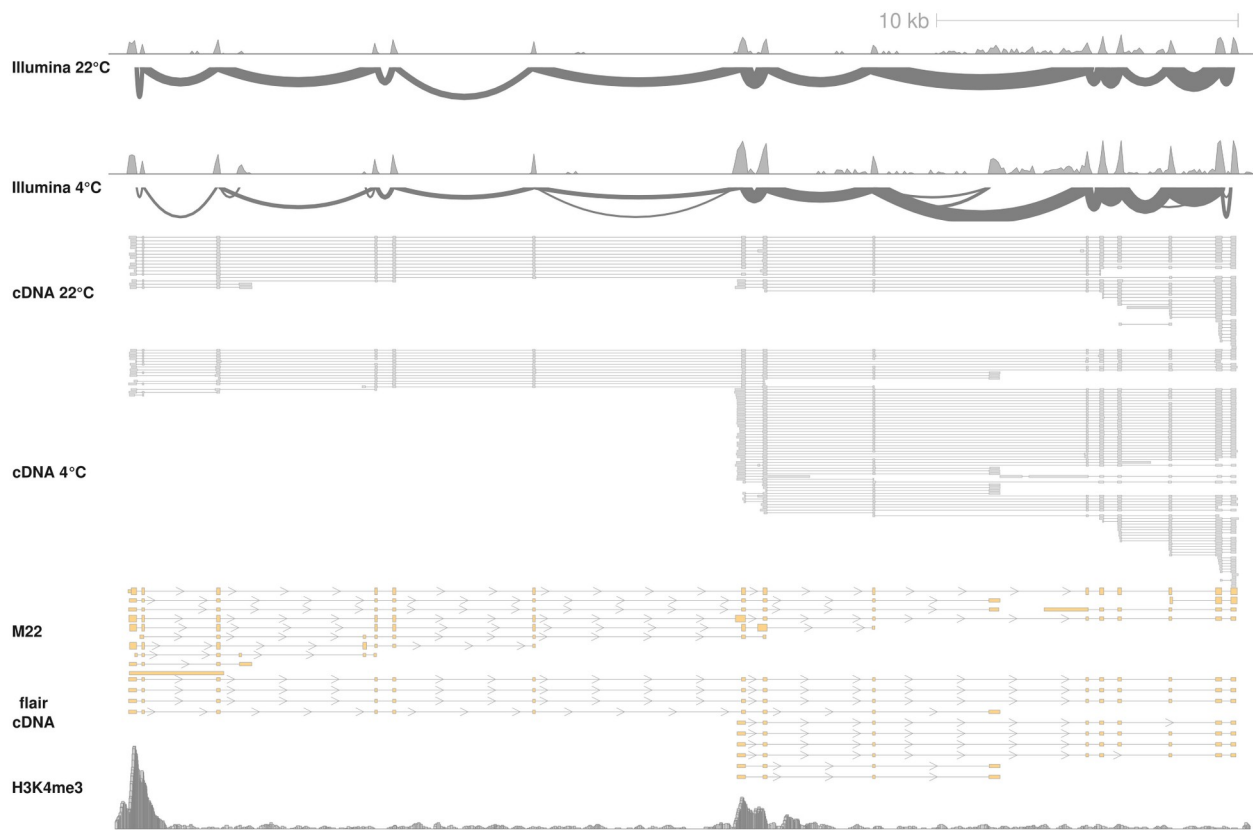
660

Fig S6: Detailed analysis of the Cars2 locus. Browser tracks of transcripts annotated for the *Cars2* locus in murine iBAT. Shown are the coverage (log10(x+1)) and splice junction usage from two representative Illumina sequencing runs, long-read alignments from two representative direct cDNA sequencing runs, as well as a the reannotated transcripts from FLAIR compared to the reference annotation (GENCODE M22)

665 and histone ChIP-Seq peaks marking the TSSs. Note the diversity in Cars2 transcripts and their usage between treatments.
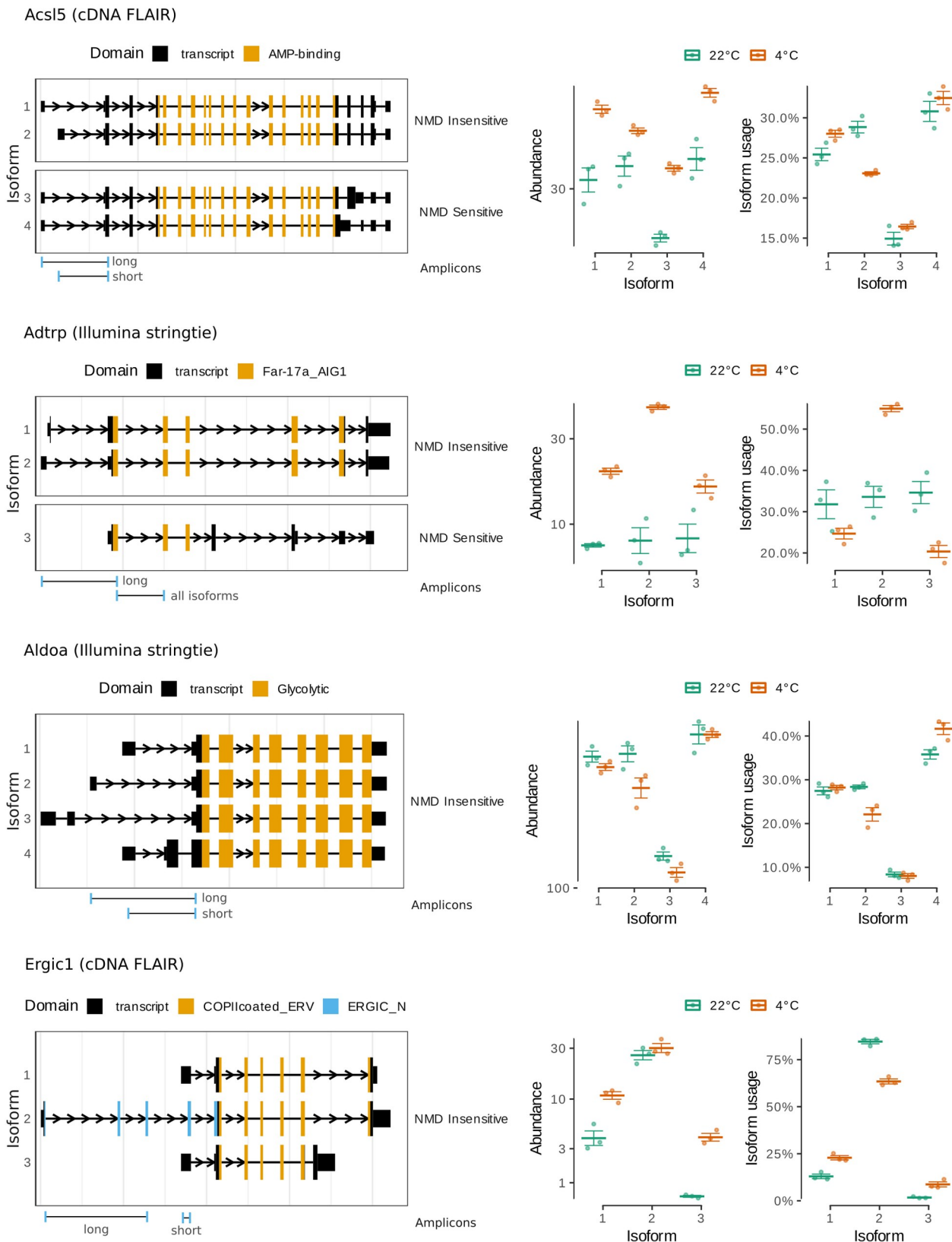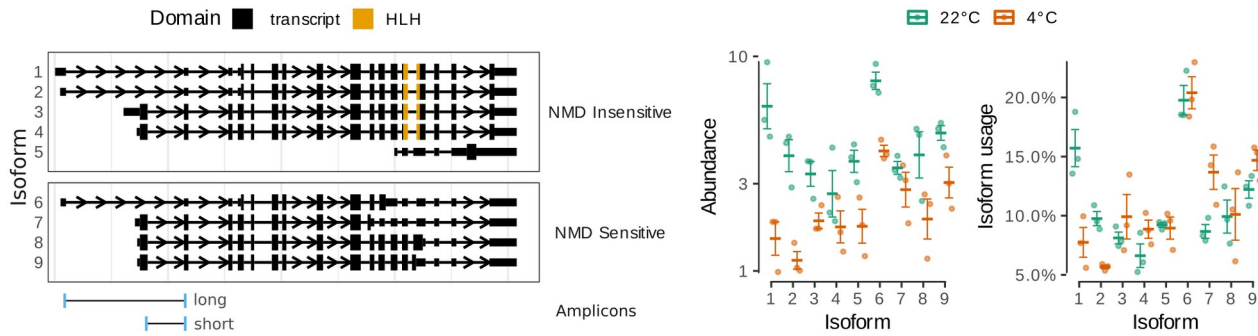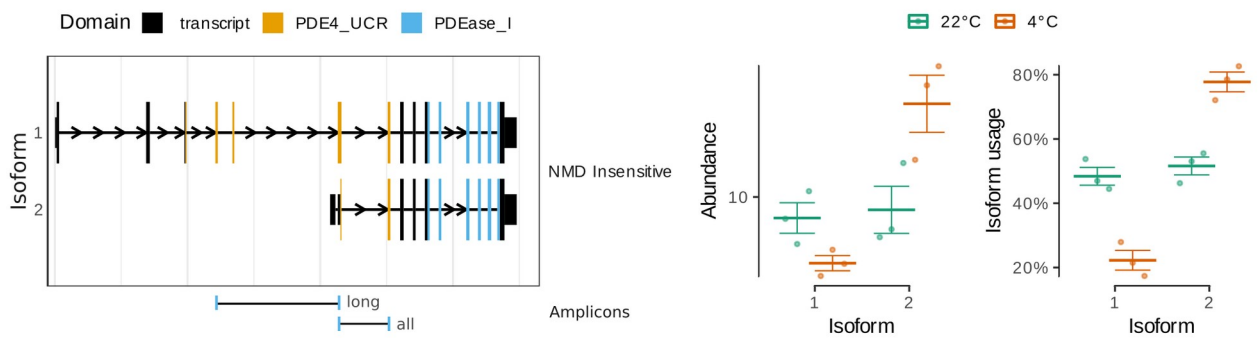
Fig S7: Differential transcript usage in cold-activated murine iBAT. Structure of expressed isoforms of genes showing differential isoform usage detected in the indicate dataset.

Mlxipl (TeloPrime stringtie)

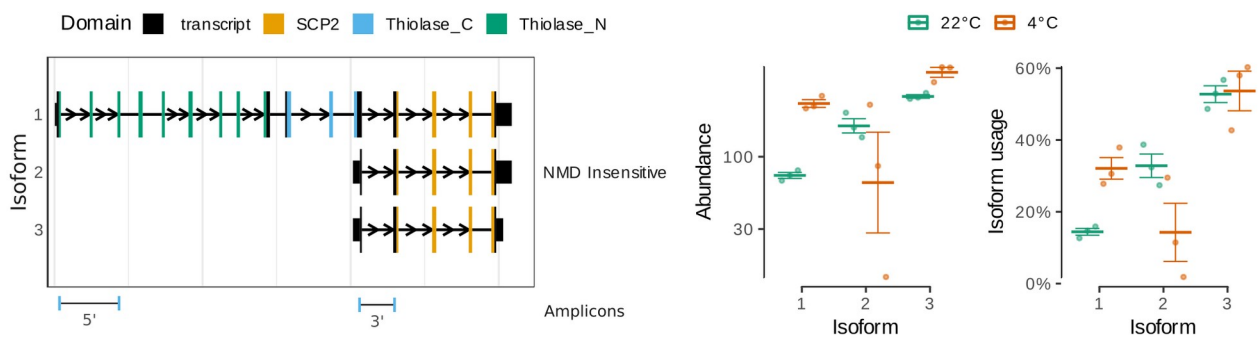## Pde4d (Illumina stringtie)

Scp2 (cDNA FLAIR)

Fig S7 (continued)

12.   Supplementary tables

Tab. S1. qPCR primers.

Target	forward	reverse
Acsl5 short	AGAAGCCCGGGGTGAGGA	AGGAAATTCAGACCCTGGGAAGG
Acsl5 long	TTCAGGTGCGCCTCGTCC	ACCAACAGGAAATTCAGACCTCTTG
Adtrp long	ACCCAGGTAGCTAAGCTTTCCA	CTGGATCGGAGCGACAGTGG
Adtrp all	TGACACTCCTTAATCTGCTCTTGC	AAGCCATTGTGGTAAAGAGCAGA
Aldoa long	TGCGCCACAGGAAAGCAC	CCTTGCCCGGAGCCACAA
Aldoa short	CTGGTGACCCTAGCCGCG	CGGTGGCAGTGCTTTCCTG
Cars2 long	CCTGAAGGTTCTGCCGCCA	ACGCTGCCTGTAGCTGTCG
Cars2 short	GCTGCACTGTGGCTCCTCA	CCGAACACCTCACTGGCCA
Ergic1 long	GCACGATGCCCTTTGACTTCA	GGCAGCAGCAGATGGAGATAA
Ergic1 short	ACCAAGCAAGGTGGCCCC	GCCCGACCAATTCACAGTGC
Mlxipl long	GCGTGGAGCTCAGTGGCA	CCAGGCTCTCCAGATGGCG
Mlxipl short	AGCCTGGCTTACAGTGGCA	CCAGGCTCTCCAGATGGCG
Pde4d all	GAGACCCTGGAGGAGCTGGA	CCTTTTGAACTTGTTGGAGGCCA
Pde4d long	TCTTGGCTAGCCTGCGAACC	GGCCTCCTCTGTGATGGTGG
Scp2 5' + full	TGGTCGGCGTTGGCATGA	ACACATAGCCAACGCATGCC
Scp2 3' + full	TGCTGGCAGCTGAGAGGC	TCCTGAAGGAGCTGGCAGC

675

13.   References

Abugessaisa, I., Noguchi, S., Hasegawa, A., Harshbarger, J., Kondo, A., Lizio, M., Severin, J., Carninci, P., Kawaji, H., & Kasukawa, T. (2017). FANTOM5 CAGE profiles of human and mouse reprocessed for GRCh38 and GRCm38 genome assemblies. *Scientific Data*, *4*(1), 170107. https://doi.org/10.1038/sdata.2017.107

Akaike, T., Ida, T., Wei, F.-Y., Nishida, M., Kumagai, Y., Alam, M. M., Ihara, H., Sawa, T., Matsunaga, T., Kasamatsu, S., Nishimura, A., Morita, M., Tomizawa, K., Nishimura, A., Watanabe, S., Inaba, K., Shima, H., Tanuma, N., Jung, M., … Motohashi, H. (2017). Cysteinyl-tRNA synthetase governs cysteine polysulfidation and mitochondrial bioenergetics. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/s41467-017-01311-y

Bartelt, A., Widenmaier, S. B., Schlein, C., Johann, K., Goncalves, R. L. S., Eguchi, K., Fischer, A. W., Parlakgül, G., Snyder, N. A., Nguyen, T. B., Bruns, O. T., Franke, D., Bawendi, M. G., Lynes, M. D., Leiria, L. O., Tseng, Y.-H., Inouye, K. E., Arruda, A. P., & Hotamisligil, G. S. (2018).

Brown adipose tissue thermogenic adaptation requires Nrf1-mediated proteasomal activity. *Nature Medicine*, *24*(3), Article 3. https://doi.org/10.1038/nm.4481

Bowman, T. A., O'Keeffe, K. R., D'Aquila, T., Yan, Q. W., Griffin, J. D., Killion, E. A., Salter, D. M., Mashek, D. G., Buhman, K. K., & Greenberg, A. S. (2016). Acyl CoA synthetase 5 (ACSL5) ablation in mice increases energy expenditure and insulin sensitivity and delays fat absorption. *Molecular Metabolism*, *5*(3), 210–220. https://doi.org/10.1016/j.molmet.2016.01.001

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525. https://doi.org/10.1038/nbt.3519

Byrne, A., Cole, C., Volden, R., & Vollmers, C. (2019). Realizing the potential of full-length transcriptome sequencing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *374*(1786), 20190097. https://doi.org/10.1098/rstb.2019.0097

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., & Rinn, J. L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development*, *25*(18), 1915–1927. https://doi.org/10.1101/gad.17446611

Cannon, B., & Nedergaard, J. (2004). Brown Adipose Tissue: Function and Physiological Significance. *Physiological Reviews*, *84*(1), 277–359. https://doi.org/10.1152/physrev.00015.2003

Chen, Y., Davidson, N. M., Wan, Y. K., Patel, H., Yao, F., Low, H. M., Hendra, C., Watten, L., Sim, A., Sawyer, C., Iakovleva, V., Lee, P. L., Xin, L., Ng, H. E. V., Loo, J. M., Ong, X., Ng, H. Q. A., Wang, J., Koh, W. Q. C., … Consortium, S.-Ne. (2021). *A systematic benchmark of Nanopore long read RNA sequencing for transcript level analysis in human cell lines* (p.

2021.04.21.440736). bioRxiv.
https://doi.org/10.1101/2021.04.21.440736

Chi, Y.-L., & Lin, J.-C. (2018). RBM4a modulates the impact of PRDM16 on development of brown adipocytes through an alternative splicing mechanism. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1865*(11, Part A), 1515–1525. https://doi.org/10.1016/j.bbamcr.2018.08.001

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, *34*(15), 2666–2669. https://doi.org/10.1093/bioinformatics/bty149

de Klerk, E., & 't Hoen, P. A. C. (2015). Alternative mRNA transcription, processing, and translation: Insights from RNA sequencing. *Trends in Genetics*, *31*(3), 128–139. https://doi.org/10.1016/j.tig.2015.01.001

De Paoli-Iseppi, R., Gleeson, J., & Clark, M. B. (2021). Isoform Age—Splice Isoform Profiling Using Long-Read Technologies. *Frontiers in Molecular Biosciences*, *8*, 711733. https://doi.org/10.3389/fmolb.2021.711733

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology*, *32*(12), 1262–1267. https://doi.org/10.1038/nbt.3026

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., deWinter, A., Dixon, J., … Turner, S. (2009). Real-Time DNA Sequencing from Single Polymerase

Molecules. *Science*, *323*(5910), 133–138. https://doi.org/10.1126/science.1162986

Engelhard, C. A., Huang, C., Khani, S., Kasparek, P., Prochazka, J., Rozman, J., Reguera, D. P., Sedlacek, R., & Kornfeld, J.-W. (2022). Comprehensive Transcriptional Profiling and Mouse Phenotyping Reveals Dispensable Role for Adipose Tissue Selective Long Noncoding RNA Gm15551. *Non-Coding RNA*, *8*(3), Article 3. https://doi.org/10.3390/ncrna8030032

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., & Bateman, A. (2016). The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Research*, *44*(D1), D279–D285. https://doi.org/10.1093/nar/gkv1344

Fiszbein, A., & Kornblihtt, A. R. (2017). Alternative splicing switches: Important players in cell differentiation. *BioEssays*, *39*(6), 1600157. https://doi.org/10.1002/bies.201600157

Floor, S. N., & Doudna, J. A. (2016). Tunable protein synthesis by transcript isoforms in human cells. *ELife*, *5*, e10921. https://doi.org/10.7554/eLife.10921

Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J. E., Mudge, J. M., Sisu, C., Wright, J. C., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Boix, C., Carbonell Sala, S., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I. T., García Girón, C., … Flicek, P. (2021). GENCODE 2021. *Nucleic Acids Research*, *49*(D1), D916–D923. https://doi.org/10.1093/nar/gkaa1087

Fu, S., Ma, Y., Yao, H., Xu, Z., Chen, S., Song, J., & Au, K. F. (2018). IDP-denovo: De novo transcriptome assembly and isoform annotation by hybrid sequencing. *Bioinformatics*, *34*(13), 2168–2176. https://doi.org/10.1093/bioinformatics/bty098

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., Jordan, M., Ciccone, J.,

Serra, S., Keenan, J., Martin, S., McNeill, L., Wallace, E. J., Jayasinghe, L., Wright, C., … Turner, D. J. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods*, *15*(3), 201–206. https://doi.org/10.1038/nmeth.4577

Gervois, P., Torra, I. P., Chinetti, G., Grötzinger, T., Dubois, G., Fruchart, J.-C., Fruchart-Najib, J., Leitersdorf, E., & Staels, B. (1999). A Truncated Human Peroxisome Proliferator-Activated Receptor α Splice Variant with Dominant Negative Activity. *Molecular Endocrinology*, *13*(9), 1535–1549. https://doi.org/10.1210/mend.13.9.0341

Glinos, D. A., Garborcauskas, G., Hoffman, P., Ehsan, N., Jiang, L., Gokden, A., Dai, X., Aguet, F., Brown, K. L., Garimella, K., Bowers, T., Costello, M., Ardlie, K., Jian, R., Tucker, N. R., Ellinor, P. T., Harrington, E. D., Tang, H., Snyder, M., … Cummings, B. B. (2022). Transcriptome variation in human tissues revealed by long-read sequencing. *Nature*, *608*(7922), Article 7922. https://doi.org/10.1038/s41586-022-05035-y

Ip, C. L. C., Loose, M., Tyson, J. R., Cesare, M. de, Brown, B. L., Jain, M., Leggett, R. M., Eccles, D. A., Zalunin, V., Urban, J. M., Piazza, P., Bowden, R. J., Paten, B., Mwaigwisya, S., Batty, E. M., Simpson, J. T., Snutch, T. P., Birney, E., Buck, D., … Consortium, M. A. and R. (2015). *MinION Analysis and Reference Consortium: Phase 1 data release and analysis* (4:1075). F1000Research. https://doi.org/10.12688/f1000research.7201.1

Joshi, A. S., Zhang, H., & Prinz, W. A. (2017). Organelle biogenesis in the endoplasmic reticulum. *Nature Cell Biology*, *19*(8), Article 8. https://doi.org/10.1038/ncb3579

Kleiboeker, B., & Lodhi, I. J. (2022). Peroxisomal regulation of energy homeostasis: Effect on obesity and related metabolic disorders. *Molecular Metabolism*, *65*, 101577. https://doi.org/10.1016/j.molmet.2022.101577

Klever, M.-K., Sträng, E., Jungnitsch, J., Melo, U. S., Hetzel, S., Dolnik, A., Schöpflin, R., Schrezenmeier, J. F., Blau, O., Westermann, J., Döhner, K., Schrezenmeier, H., Spielmann, M., Meissner, A., Mundlos, S., & Bullinger, L. (2020). Integration of Hi-C and Nanopore Sequencing for Structural Variant Analysis in AML with a Complex Karyotype: (Chromothripsis)[2]. *Blood*, *136*, 28. https://doi.org/10.1182/blood-2020-133787

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O., & Zhang, F. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, *517*(7536), 583–588. https://doi.org/10.1038/nature14136

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology*, *20*(1), 278. https://doi.org/10.1186/s13059-019-1910-1

Lau, E., Han, Y., Williams, D. R., Thomas, C. T., Shrestha, R., Wu, J. C., & Lam, M. P. Y. (2019). Splice-Junction-Based Mapping of Alternative Isoforms in the Human Proteome. *Cell Reports*, *29*(11), 3751-3765.e5. https://doi.org/10.1016/j.celrep.2019.11.026

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, *9*(8), e1003118. https://doi.org/10.1371/journal.pcbi.1003118

Leshkowitz, D., Kedmi, M., Fried, Y., Pilzer, D., Keren-Shaul, H., Ainbinder, E., & Dassa, B. (2022). Exploring differential exon usage via short- and long-read RNA sequencing strategies. *Open Biology*, *12*(9), 220206. https://doi.org/10.1098/rsob.220206

Leung, S. K., Jeffries, A. R., Castanho, I., Jordan, B. T., Moore, K., Davies, J. P., Dempster, E. L., Bray, N. J., O'Neill, P., Tseng, E., Ahmed, Z., Collier, D.

A., Jeffery, E. D., Prabhakar, S., Schalkwyk, L., Jops, C., Gandal, M. J., Sheynkman, G. M., Hannon, E., & Mill, J. (2021). Full-length transcript sequencing of human and mouse cerebral cortex identifies widespread isoform diversity and alternative splicing. *Cell Reports*, *37*(7). https://doi.org/10.1016/j.celrep.2021.110022

Li, D., Zhang, F., Zhang, X., Xue, C., Namwanje, M., Fan, L., Reilly, M. P., Hu, F., & Qiang, L. (2016). Distinct functions of PPARγ isoforms in regulating adipocyte plasticity. *Biochemical and Biophysical Research Communications*, *481*(1–2), 132–138. https://doi.org/10.1016/j.bbrc.2016.10.152

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, N. C., Fan, J., & Papadopoulos, V. (2016). Sterol Carrier Protein-2, a Nonspecific Lipid-Transfer Protein, in Intracellular Cholesterol Trafficking in Testicular Leydig Cells. *PLOS ONE*, *11*(2), e0149728. https://doi.org/10.1371/journal.pone.0149728

Li, P., Song, R., Du, Y., Liu, H., & Li, X. (2022). Adtrp regulates thermogenic activity of adipose tissue via mediating the secretion of S100b. *Cellular and Molecular Life Sciences: CMLS*, *79*(8), 407. https://doi.org/10.1007/s00018-022-04441-9

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Love, M. I., Soneson, C., & Patro, R. (2018). Swimming downstream: Statistical analysis of differential transcript usage following Salmon quantification. *F1000Research*, *7*, 952. https://doi.org/10.12688/f1000research.15398.3

Lundh, M., Pluciñska, K., Isidor, M. S., Petersen, P. S. S., & Emanuelli, B. (2017). Bidirectional manipulation of gene expression in adipocytes

using CRISPRa and siRNA. *Molecular Metabolism*, *6*(10), 1313–1320. https://doi.org/10.1016/j.molmet.2017.07.001

Lynch, M. J., Baillie, G. S., & Houslay, M. D. (2007). CAMP-specific phosphodiesterase-4D5 (PDE4D5) provides a paradigm for understanding the unique non-redundant roles that PDE4 isoforms play in shaping compartmentalized cAMP cell signalling. *Biochemical Society Transactions*, *35*(5), 938–941. https://doi.org/10.1042/BST0350938

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), Article 1. https://doi.org/10.14806/ej.17.1.200

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, *19*(6), Article 6. https://doi.org/10.1038/s41592-022-01488-1

Moura, M. A. F., Festuccia, W. T. L., Kawashita, N. H., Garófalo, M. A. R., Brito, S. R. C., Kettelhut, I. C., & Migliorini, R. H. (2005). Brown adipose tissue glyceroneogenesis is activated in rats exposed to cold. *Pflügers Archiv*, *449*(5), 463–469. https://doi.org/10.1007/s00424-004-1353-7

Nowicka, M., & Robinson, M. D. (2016). DRIMSeq: A Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Research*, *5*, 1356. https://doi.org/10.12688/f1000research.8900.2

Oikonomopoulos, S., Bayega, A., Fahiminiya, S., Djambazian, H., Berube, P., & Ragoussis, J. (2020). Methodologies for Transcript Profiling Using Long-Read Technologies. *Frontiers in Genetics*, *11*. https://www.frontiersin.org/articles/10.3389/fgene.2020.00606

Oikonomopoulos, S., Wang, Y. C., Djambazian, H., Badescu, D., & Ragoussis, J. (2016). Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Scientific Reports*, *6*(1), 31602. https://doi.org/10.1038/srep31602

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome

by high-throughput sequencing. *Nature Genetics*, *40*(12), Article 12. https://doi.org/10.1038/ng.259

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Rajendran, V., Kalita, P., Shukla, H., Kumar, A., & Tripathi, T. (2018). Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. *International Journal of Biological Macromolecules*, *111*, 400–414. https://doi.org/10.1016/j.ijbiomac.2017.12.157

Reverte-Salisa, L., Sanyal, A., & Pfeifer, A. (2019). Role of cAMP and cGMP Signaling in Brown Fat. *Handbook of Experimental Pharmacology*, *251*, 161–182. https://doi.org/10.1007/164_2018_117

Ringeling, F. R., Chakraborty, S., Vissers, C., Reiman, D., Patel, A. M., Lee, K.-H., Hong, A., Park, C.-W., Reska, T., Gagneur, J., Chang, H., Spletter, M. L., Yoon, K.-J., Ming, G., Song, H., & Canzar, S. (2022). Partitioning RNAs by length improves transcriptome reconstruction from short-read RNA-seq data. *Nature Biotechnology*, *40*(5), Article 5. https://doi.org/10.1038/s41587-021-01136-7

Robinson, E. K., Jagannatha, P., Covarrubias, S., Cattle, M., Smaliy, V., Safavi, R., Shapleigh, B., Abu-Shumays, R., Jain, M., Cloonan, S. M., Akeson, M., Brooks, A. N., & Carpenter, S. (2021). Inflammation drives alternative first exon usage to regulate immune genes including a novel iron-regulated isoform of Aim2. *ELife*, *10*, e69431. https://doi.org/10.7554/eLife.69431

Rodriguez, J. M., Pozo, F., Domenico, T. di, Vazquez, J., & Tress, M. L. (2020). An analysis of tissue-specific alternative splicing at the protein level. *PLOS Computational Biology*, *16*(10), e1008287. https://doi.org/10.1371/journal.pcbi.1008287

Rosen, E. D., & Spiegelman, B. M. (2014). What We Talk About When We Talk About Fat. *Cell*, *156*(1), 20–44. https://doi.org/10.1016/j.cell.2013.12.012

Seale, P. (2015). Transcriptional Regulatory Circuits Controlling Brown Fat Development and Activation. *Diabetes*, *64*(7), 2369–2375. https://doi.org/10.2337/db15-0203

Sessegolo, C., Cruaud, C., Silva, C. D., Cologne, A., Dubarry, M., Derrien, T., Lacroix, V., & Aury, J.-M. (2019). Transcriptome profiling of mouse samples using nanopore sequencing of cDNA and RNA molecules. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-51470-9

Shapira, S. N., & Seale, P. (2019). Transcriptional Control of Brown and Beige Fat Development and Function. *Obesity*, *27*(1), 13–21. https://doi.org/10.1002/oby.22334

Sharon, D., Tilgner, H., Grubert, F., & Snyder, M. (2013). A single-molecule long-read survey of the human transcriptome. *Nature Biotechnology*, *31*(11), 1009–1014. https://doi.org/10.1038/nbt.2705

Shumate, A., Wong, B., Pertea, G., & Pertea, M. (2021). *Improved Transcriptome Assembly Using a Hybrid of Long and Short Reads with StringTie* (p. 2021.12.08.471868). https://doi.org/10.1101/2021.12.08.471868

Soneson, C., Love, M. I., & Robinson, M. D. (2016). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research*, *4*, 1521. https://doi.org/10.12688/f1000research.7563.2

Soneson, C., Yao, Y., Bratus-Neuenschwander, A., Patrignani, A., Robinson, M. D., & Hussain, S. (2019). A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes.

Nature Communications, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-11272-z

Soriano, R. N., Braga, S. P., Breder, J. S. C., Batalhao, M. E., Oliveira-Pelegrin, G. R., Ferreira, L. F. R., Rocha, M. J. A., Carnio, E. C., & Branco, L. G. S. (2018). Endogenous peripheral hydrogen sulfide is propyretic: Its permissive role in brown adipose tissue thermogenesis in rats. *Experimental Physiology*, *103*(3), 397–407. https://doi.org/10.1113/EP086775

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: The teenage years. *Nature Reviews Genetics*, *20*(11), 631–656. https://doi.org/10.1038/s41576-019-0150-2

Stephens, M. (2017). False discovery rates: A new deal. *Biostatistics*, *18*(2), 275–294. https://doi.org/10.1093/biostatistics/kxw041

Tang, A. D., Soulette, C. M., van Baren, M. J., Hart, K., Hrabeta-Robinson, E., Wu, C. J., & Brooks, A. N. (2020). Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-15171-6

Tardaguila, M., Fuente, L. de la, Marti, C., Pereira, C., Pardo-Palacios, F. J., Risco, H. del, Ferrell, M., Mellado, M., Macchietto, M., Verheggen, K., Edelmann, M., Ezkurdia, I., Vazquez, J., Tress, M., Mortazavi, A., Martens, L., Rodriguez-Navarro, S., Moreno-Manzano, V., & Conesa, A. (2018). SQANTI: Extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. *Genome Research*, *28*(3), 396–411. https://doi.org/10.1101/gr.222976.117

Tomita, H., Shakkottai, V. G., Gutman, G. A., Sun, G., Bunney, W. E., Cahalan, M. D., Chandy, K. G., & Gargus, J. J. (2003). Novel truncated isoform of SK3 potassium channel is a potent dominant-negative regulator of SK

currents: Implications in schizophrenia. *Molecular Psychiatry*, *8*(5), 524–535. https://doi.org/10.1038/sj.mp.4001271

Udaondo, Z., Sittikankaew, K., Uengwetwanit, T., Wongsurawat, T., Sonthirod, C., Jenjaroenpun, P., Pootakham, W., Karoonuthaisiri, N., & Nookaew, I. (2021). Comparative Analysis of PacBio and Oxford Nanopore Sequencing Technologies for Transcriptomic Landscape Identification of Penaeus monodon. *Life (Basel, Switzerland)*, *11*(8), 862. https://doi.org/10.3390/life11080862

Van den Berge, K., Soneson, C., Robinson, M. D., & Clement, L. (2017). stageR: A general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. *Genome Biology*, *18*(1), 151. https://doi.org/10.1186/s13059-017-1277-0

Vernia, S., Edwards, Y. J., Han, M. S., Cavanagh-Kyros, J., Barrett, T., Kim, J. K., & Davis, R. J. (2016). An alternative splicing program promotes adipose tissue thermogenesis. *ELife*, *5*, e17672. https://doi.org/10.7554/eLife.17672

Vezzosi, D., & Bertherat, J. (2011). Phosphodiesterases in endocrine physiology and disease. *European Journal of Endocrinology*, *165*(2), 177–188. https://doi.org/10.1530/EJE-10-1123

Vitting-Seerup, K., Porse, B. T., Sandelin, A., & Waage, J. (2014). spliceR: An R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics*, *15*(1), 81. https://doi.org/10.1186/1471-2105-15-81

Vitting-Seerup, K., & Sandelin, A. (2019). IsoformSwitchAnalyzeR: Analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics*, *35*(21), 4469–4471. https://doi.org/10.1093/bioinformatics/btz247

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., & Burge, C. B. (2008). Alternative

isoform regulation in human tissue transcriptomes. *Nature*, *456*(7221), 470–476. https://doi.org/10.1038/nature07509

Wang, W., & Seale, P. (2016). Control of brown and beige fat development. *Nature Reviews Molecular Cell Biology*, *17*(11), nrm.2016.96. https://doi.org/10.1038/nrm.2016.96

Weatheritt, R. J., Sterne-Weiler, T., & Blencowe, B. J. (2016). The ribosome-engaged landscape of alternative splicing. *Nature Structural & Molecular Biology*, *23*(12), 1117–1123. https://doi.org/10.1038/nsmb.3317

Wei, C., Ma, X., Su, K., Qi, S., Zhu, Y., Lin, J., Wang, C., Yang, R., Chen, X., Wang, W., & Zhang, W. J. (2020). ChREBP-β regulates thermogenesis in brown adipose tissue. *Journal of Endocrinology*, *245*(3), 343–356. https://doi.org/10.1530/JOE-19-0498

Workman, R. E., Tang, A. D., Tang, P. S., Jain, M., Tyson, J. R., Razaghi, R., Zuzarte, P. C., Gilpatrick, T., Payne, A., Quick, J., Sadowski, N., Holmes, N., Jesus, J. G. de, Jones, K. L., Soulette, C. M., Snutch, T. P., Loman, N., Paten, B., Loose, M., … Timp, W. (2019). Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nature Methods*, *16*(12), 1297–1305. https://doi.org/10.1038/s41592-019-0617-2

Wright, D. J., Hall, N. A. L., Irish, N., Man, A. L., Glynn, W., Mould, A., Angeles, A. D. L., Angiolini, E., Swarbreck, D., Gharbi, K., Tunbridge, E. M., & Haerty, W. (2022). Long read sequencing reveals novel isoforms and insights into splicing regulation during cell state changes. *BMC Genomics*, *23*(1), 42. https://doi.org/10.1186/s12864-021-08261-2

Xu, Q., Modrek, B., & Lee, C. (2002). Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Research*, *30*(17), 3754–3766. https://doi.org/10.1093/nar/gkf492

Yi, X., Yang, Y., Wu, P., Xu, X., & Li, W. (2020). Alternative splicing events during adipogenesis from hMSCs. *Journal of Cellular Physiology*, *235*(1), 304–316. https://doi.org/10.1002/jcp.28970

Zhu, A., Ibrahim, J. G., & Love, M. I. (2018). Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/bty895