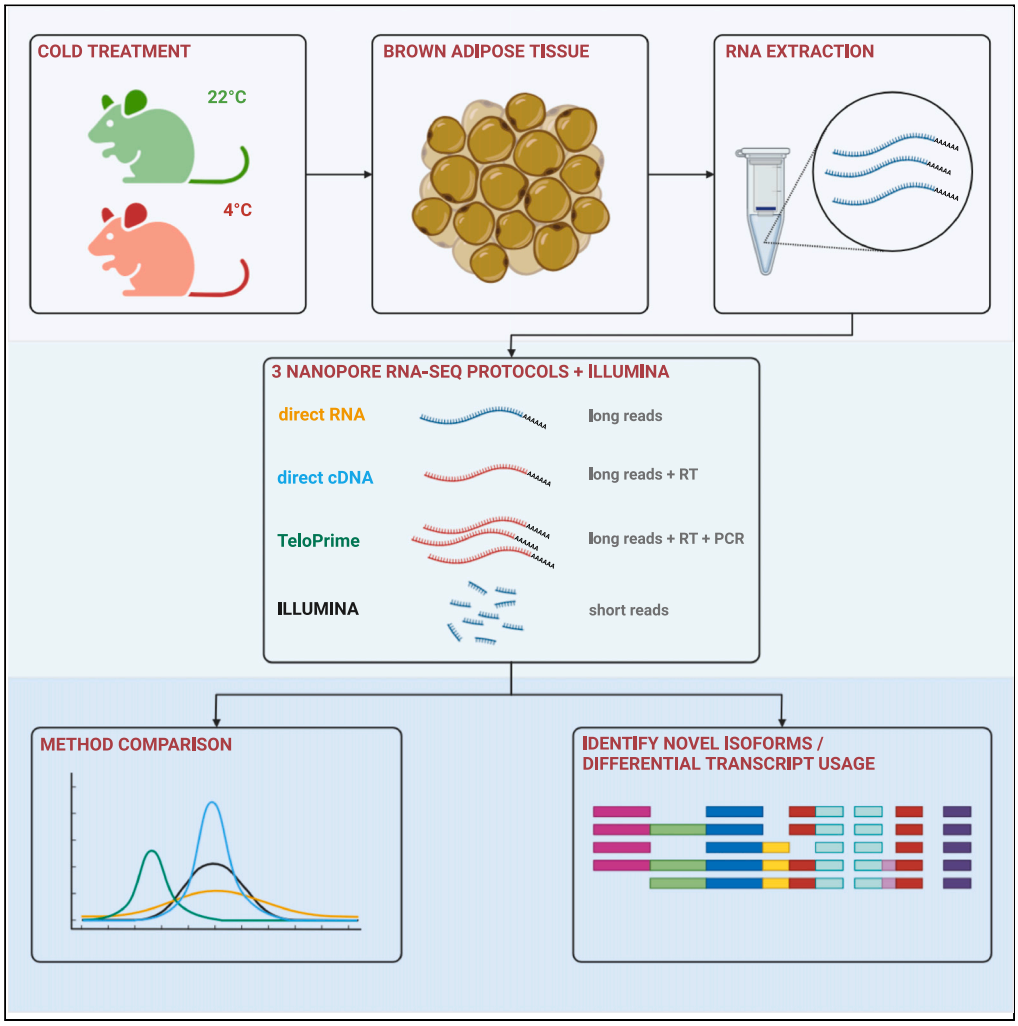
iScience

CellPress
OPEN ACCESS

Article

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




Christoph Andreas Engelhard, Sajjad Khani, Sophia Derdak, Martin Bilban, Jan-Wilhelm Kornfeld

martin.bilban@meduniwien.ac.at (M.B.)
janwilhelmkornfeld@bmb.sdu.dk (J.-W.K.)

## Highlights

Benchmark of three nanopore RNA sequencing protocols

Reannotation of murine brown adipose tissue transcriptome

Identification of differential transcript usage in murine brown adipose tissue

Engelhard et al., iScience *26*, 107190
August 18, 2023 
https://doi.org/10.1016/j.isci.2023.107190

iScience

CellPress
OPEN ACCESS

Article

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

Christoph Andreas Engelhard,[1] Sajjad Khani,[2,3] Sophia Derdak,[4] Martin Bilban,[5,6,*] and Jan-Wilhelm Kornfeld[1,6,7,*]

## SUMMARY

**Alternative transcription increases transcriptome complexity by expression of multiple transcripts per gene. Annotation and quantification of transcripts using short-read sequencing is non-trivial. Long-read sequencing aims at overcoming these problems by sequencing full-length transcripts. Activation of brown adipose tissue (BAT) thermogenesis involves major transcriptomic remodeling and positively affects metabolism via increased energy expenditure. We benchmark Oxford Nanopore Technology (ONT) long-read sequencing protocols to Illumina short-read sequencing assessing alignment characteristics, gene and transcript detection and quantification, differential gene and transcript expression, transcriptome reannotation, and differential transcript usage (DTU). We find ONT sequencing is superior to Illumina for transcriptome reassembly, reducing the risk of false-positive events by unambiguously mapping reads to transcripts. We identified novel isoforms of genes undergoing DTU in cold-activated BAT including *Cars2*, *Adtrp*, *Acsl5*, *Scp2*, *Aldoa*, and *Pde4d*, validated by real-time PCR. The reannotated murine BAT transcriptome established here provides a framework for future investigations into the regulation of BAT.**

## INTRODUCTION

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 translational complexity in a cell.[1] Alternative transcription is estimated to occur within 92%–94% of human genes, substantially expanding the catalog of co-expressed mRNAs.[2,3] 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.[4–7] Interestingly, AT is tissue specific[6,8] or marks specific cellular states,[9,10] 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 densely packed with mitochondria and morphologically present multiple small lipid droplets.[11] 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 adipocytes 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.[12] 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 crucial for e.g., the regulation of adipocyte thermogenesis.[13] Transcriptomic studies have shown that DTU is required for adipogenesis, the process of differentiation of preadipocytes into mature adipocytes.[9,14] 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 adipocyte function.[15,16]

[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 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

[6]These authors contributed equally

[7]Lead contact

*Correspondence: martin.bilban@meduniwien.ac.at (M.B.), janwilhelmkornfeld@bmb.sdu.dk (J.-W.K.)

https://doi.org/10.1016/j.isci.2023.107190

However, most studies focusing 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 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.[17] Short-read sequencing also suffers in respect to transcript level quantification required for analysis of DTU, as only reads mapping to parts 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.[18,19] Long-read sequencing methods such as those developed by Pacific Biosciences[20,21] and Oxford Nanopore Technologies (ONT)[22] generate full-length isoform reads that mitigate these limitations, allowing for simple transcriptome reannotation and unambiguous read assignment.[23] 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.[24] Reference databases like GENCODE are based on a limited number of tissue transcriptomes.[8,10] 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 other hand, suffer from lower throughput and lower base calling accuracy, resulting in failure to detect lowly expressed isoforms and fuzzy splice junction annotation.[23] Accordingly, algorithms that combine short and long reads for improved transcriptome reassembly, either *de novo* or using a reference genome, have been developed[25–27] and consortia such as GENCODE have started incorporating long-read sequencing in their reference transcriptome annotation pipelines.[28]

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

## 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[23,29] 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 an 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 (Figure 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 $13.5 \times 10^6$ reads from TeloPrime sequencing, $12{,}7 \times 10^6$

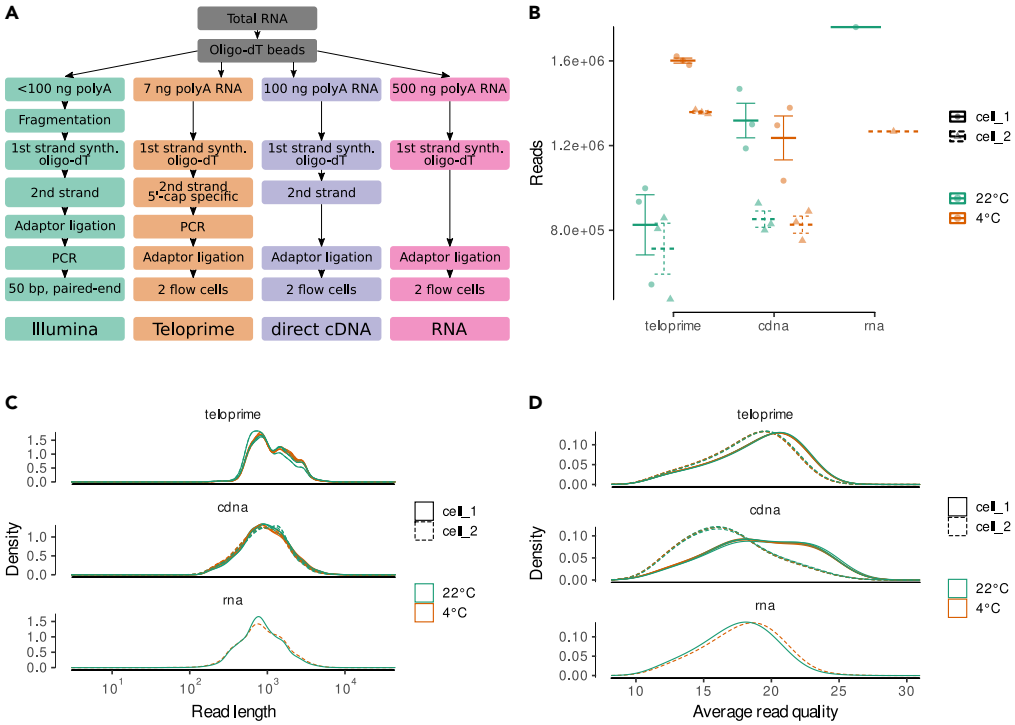



**Figure 1. Characterization of ONT reads**

(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 STAR Methods for details.

(B–D) Total number of reads (B), read length distribution (C), and read quality distribution (D) by ONT sequencing method, flow cell and housing temperature. Data are represented as mean ± SEM.

reads from direct cDNA sequencing, and $3{,}0 \times 10^6$ from the direct RNA sequencing (Table 1). Interestingly, we observed a high variability in the number of high-quality reads produced per flow cell, 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 (Figure 1B). Similar variability in MinION flow cell performance has been noticed in other studies.[30–32] 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 (Figure 1C). In good agreement with other reports,[33] 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 (Figure 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.

## 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 (Table 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 (Table 1; Figures 2A and S1A–S1C). Reads with supplementary alignments were most common in direct cDNA-Seq and these were longer (Figures 2C and S1E) and showed larger unaligned parts (Figures 2D and S1F) then reads from the other methods. Supplementary alignments can arise from reads mapping to different chromosomes, indicative for chromosomal rearrangements.[34,35] However, in direct cDNA-Seq, supplementary alignments mostly mapped

**Table 1. Alignment characteristics**

| 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 |

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)

anti-sense to the same transcript as the primary alignment (Figures 2B and 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 toward shorter reads (Figure S1G). This was less prominent for the TeloPrime and directs 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[36–38] (Figure 2E). All long-read libraries showed decreasing coverage from the 3′ end of transcripts toward 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[37] (Figures 2F and S1H). This could have been caused by RNA degradation during library preparation protocols or software artifacts during the base-calling process.[26,32] Overall, coverage was higher in TeloPrime-Seq as reported.[32] 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.

## 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,[32] direct cDNA and RNA-Seq detected a similar number of features as short-read sequencing at any given sequencing depth far outperforming TeloPrime-Seq (Figures 3A and 3B). 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 (Figure 3C). To detect the cause of these differences, we stratified the transcript detection rates by transcript length and transcript biotype (Figure 3D). As reported, detection rates for Illumina, direct cDNA-Seq and direct RNA-Seq increased with transcript size.[37] 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 (Figures S2A and S2B). Since coding genes are generally higher expressed than other classes of RNA,[39] we assessed a potential effect of expression level on the gene and transcript detection rates by the ONT methods (Figures 3E and S2C).

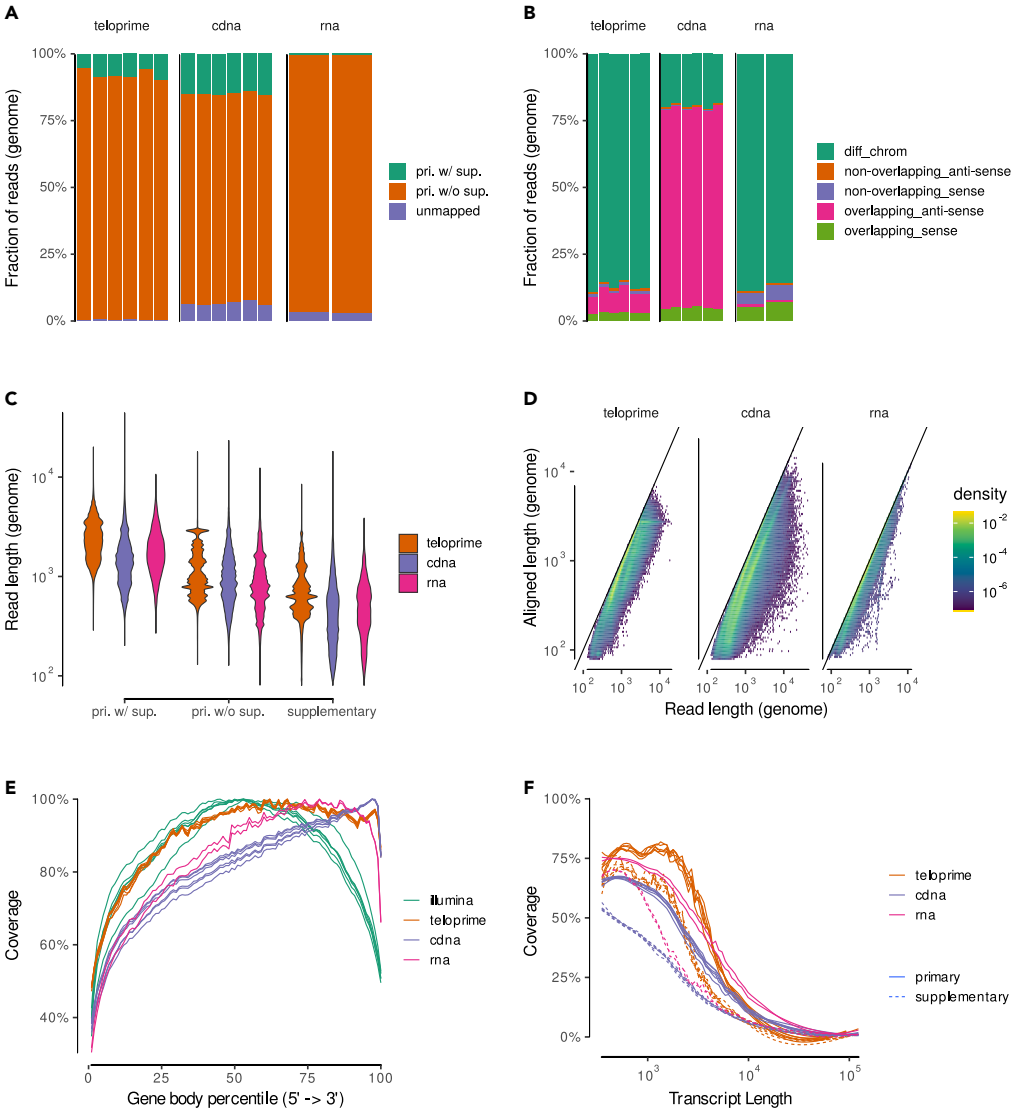



**Figure 2. Characterization of read alignments**

(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. See also Figure S1.

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 (Figures S2D and S2E). Interestingly, there were also features with high expression in the ONT datasets that were 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, while TeloPrime yielded lower detection rates with a non-linear relationship to transcript length.

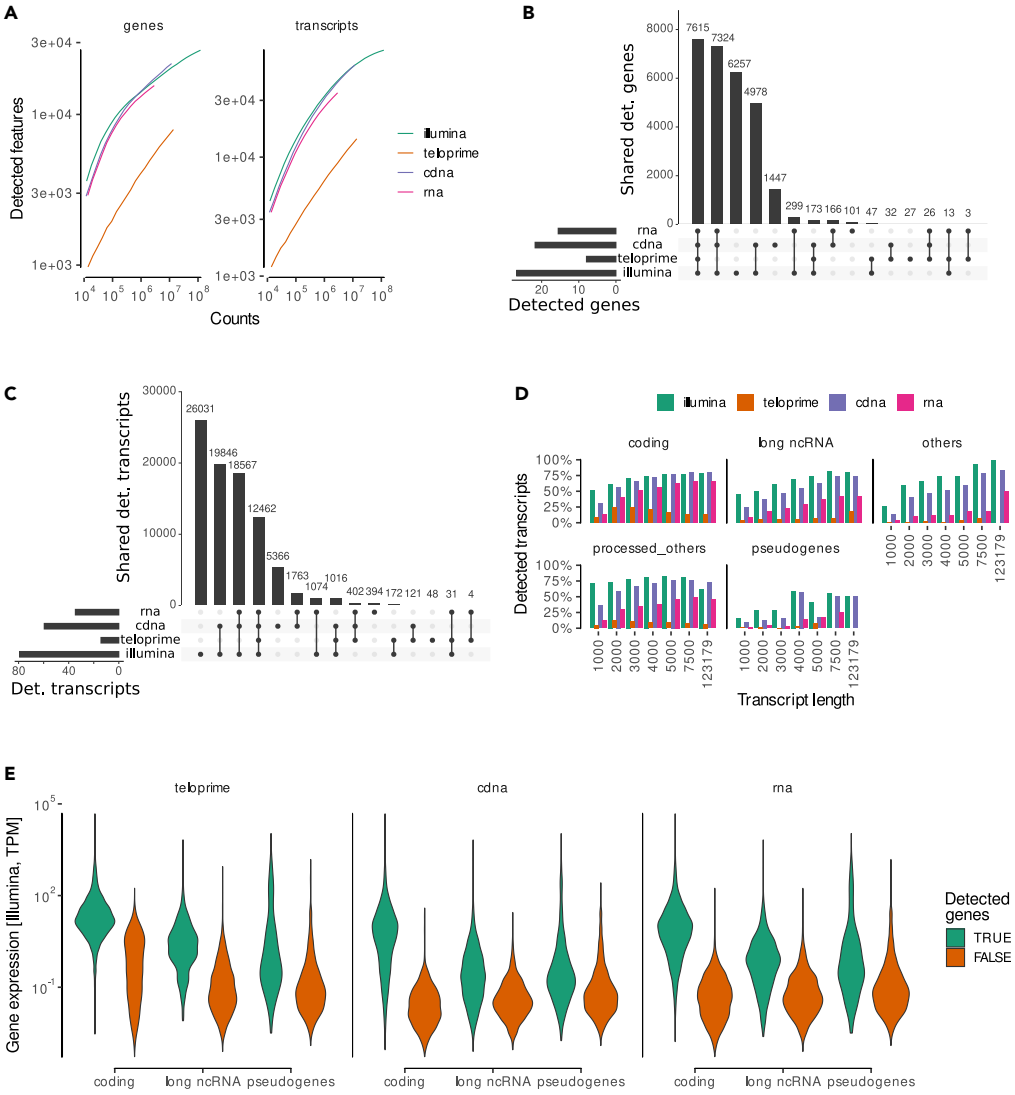



**Figure 3. Feature detection**

(A) Feature detection rate by library size. A feature is counted as detected if there is at least one primary alignment to it.

(B and 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. See also Figure S2.

## 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; Figure 4) and transcript level ($R^2$ = 0.92; Figure 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.[40] 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.[29] As we sequenced all TeloPrime and direct cDNA samples on two

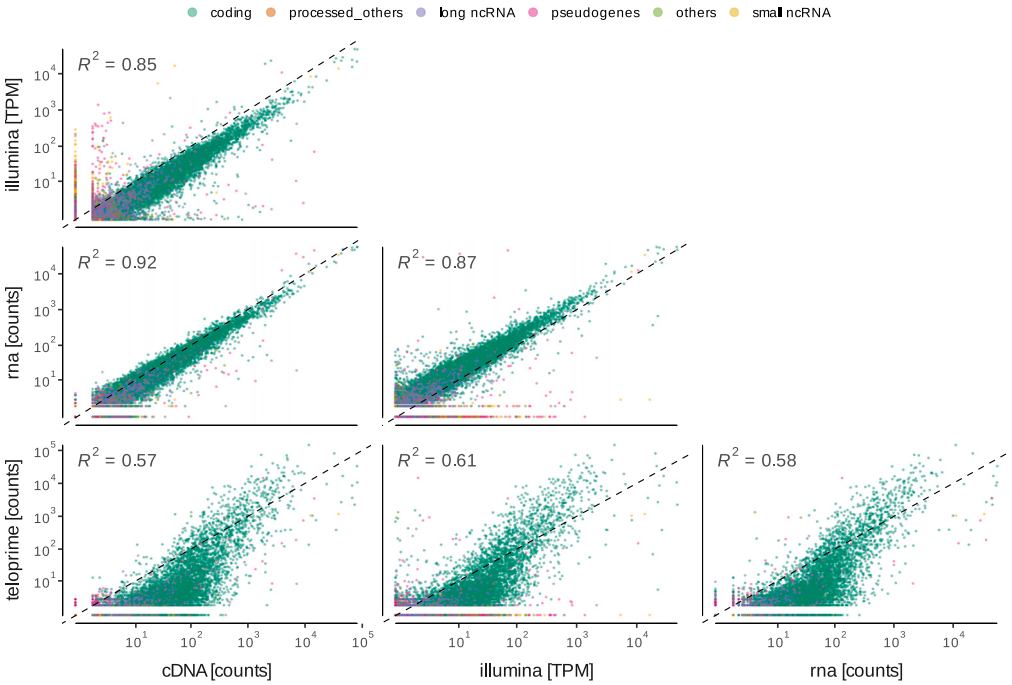



**Figure 4. Gene quantification**

Scatterplots showing the correlation in gene quantification between the different sequencing methods. See also Figure S3.

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; Figure S3B), the direct cDNA-Seq method showed a higher variation ($R^2$ = 0.75 to 0.77; Figure S3C) reflecting the variability in read quality and length distributions (Figures 1B–1D). 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.

## 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) (Figure 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 (Figure 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 (Figure 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 (Figure S4).

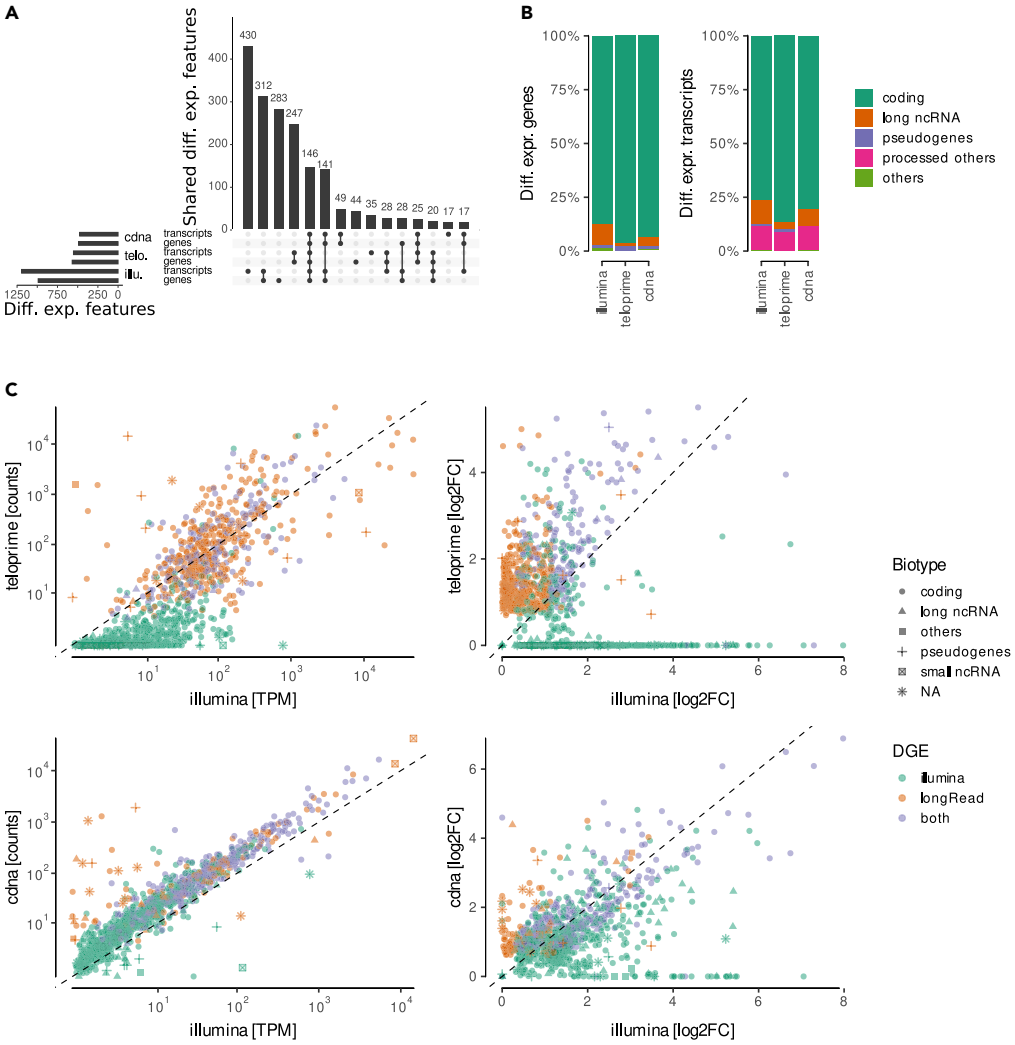



**Figure 5. Differential gene expression analysis**

(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. See also Figure S4.

## 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 BAT, we applied two transcriptome reassembly algorithms. FLAIR[27] 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[41,42] 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 (Figure 6A). Generally, StringTie reannotated more transcripts compared to FLAIR (Figure 6A), as shown previously.[26] Irrespective of the reannotation algorithm, direct cDNA sequencing yielded the highest, while direct RNA sequencing gave the lowest number of reannotated transcripts, resembling the

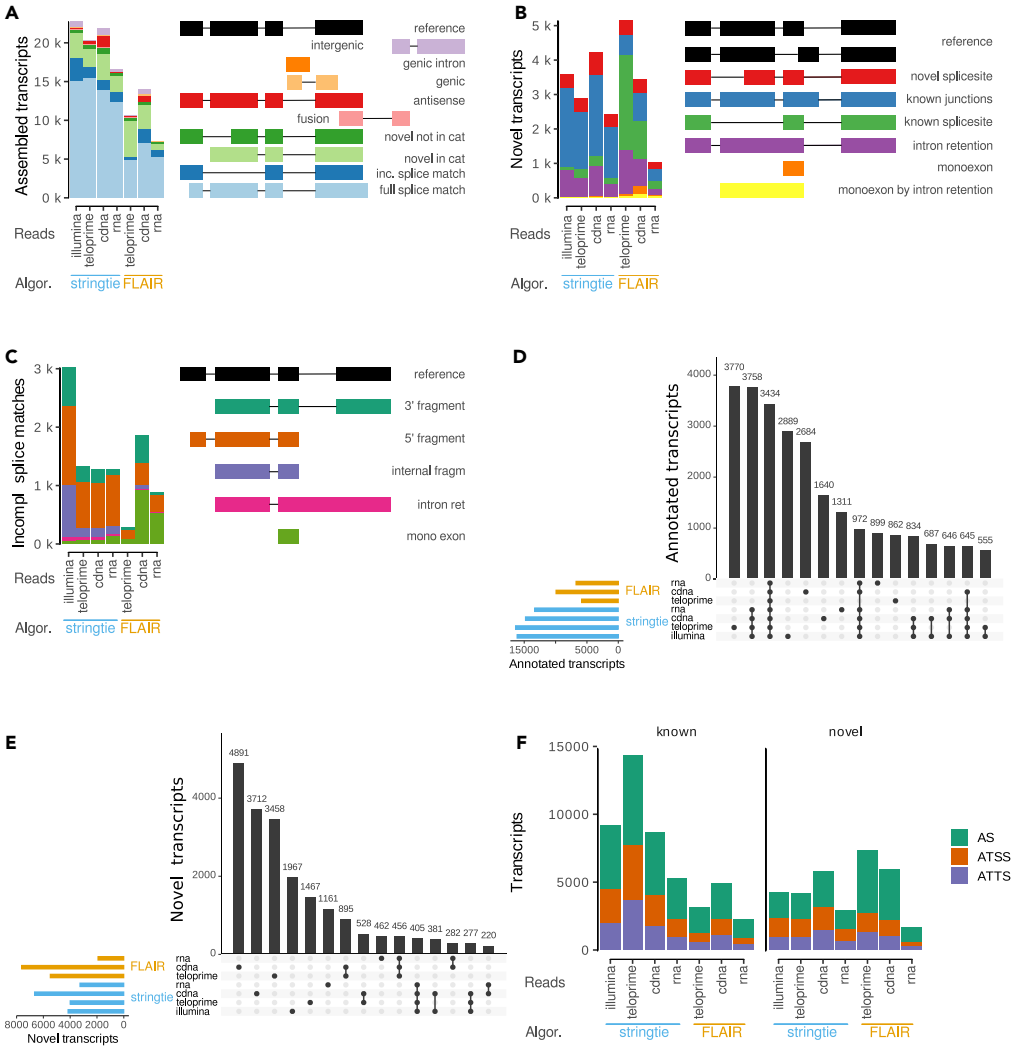



**Figure 6. Reannotation analysis of short- and long-read sequencing protocols stratified by transcript assembler**

(A) Overview of different structural categories.

(B and C) Sub-classification within the novel transcript and incomplete splice match categories.

(D) Overlap between transcripts in the reference annotation (gffcompare =, c and m) annotated by the different transcriptome assemblers and sequencing methods.

(E) Same as D but for novel transcripts not in the reference.

(F) Splice type compared to the hypothetical longest pre-mRNA for known and novel transcripts. See also Figure S5.

number of mapped reads (Figures 3B–3D), 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"; Figure 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" (Figure 6B). Among the ISM, class 5′ fragments as well as "mono-exonsmatches," were most often found (Figure 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 (Figure 6A). These were on average shorter (Figure S5A) and consisted of less exons compared to the full splice matching transcripts (Figure S5B), indicating they might be artifacts. 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 large 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 (Figure 6D). Intriguingly, large sets of annotated transcripts were only detected in the TeloPrime/StringTie, cDNA/FLAIR or in the Illumina dataset 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. Only a minority of novel transcript were identified in more than one dataset and interestingly only within the same algorithm (Figure 6E). In both known and novel transcripts, alternative splicing was more common than alternative transcription start and termination sites (Figure 6F). Thus, the ONT sequencing method and the algorithm used for reannotation had a significant impact not only on the number but also on the nature of the structural category of novel transcripts identified.

## 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 TSS or alternative splicing.[43] To overcome this limitation, we applied 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 direct cDNA counts) and reannotation algorithms (Figure 7A). As observed for the differential gene expression analysis (Figure 5A), a higher number of DTUs were identified when Illumina counts were used to assess transcript quantification (Figure 7A) and alternative splicing was the most common mechanism driving DTU, before alternative transcription start and termination sites (Figure S6B). We found little overlap between StringTie and FLAIR reannotations in line with other reports.[44] Among the genes with significant isoform switches between cold-activated BAT compared to the controls, we observed phosphodiesterase 4D (*Pde4d*), regulating levels of the signaling intermediate cAMP, which activates lipolysis, glucose uptake, and thermogenesis in brown adipocytes;[45] the thermogenesis regulating hydrolase androgen dependent TFPI regulating protein (*Adtrp)46* and regulators of fatty acid metabolism (acyl-CoA synthetase long-chain family member 5, *Acsl5;* Perilipin 5, *Plin5*), glycolysis (Aldolase A; *Aldoa*), protein sorting (endoplasmic reticulum-golgi intermediate compartment 1; *Ergic*1; Myosin light polypeptide 6; *Myl6*), lipid synthesis (MLX interacting protein-like, *Mlxipl;* choline phospotransferase 1, *Chpt1*), beta-oxidation (sterol carrier protein 2, liver, *Scp2;* Enoyl-CoA Delta Isomerase 2, *Eci2*) and protein cysteinylation (cysteinyl-tRNA synthetase 2; *Cars247;* Figures 7C–7E, S7, and S9). qPCR analysis on iBAT from control and cold-treated mice using primer sets specific for the individual transcripts corroborated our DTU analysis, thus validating the isoform regulation (Figure 7B). Incorporation of previously published histone modification chromatin immunoprecipitation (ChIP)-Seq peaks[46] 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; Figure S6A). Visual inspection of ONT direct cDNA sequencing coverage data indicated that the canonical TSS is dominant at room temperature, while 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,[47] but additionally executes a "non-canonical" function in post-translational cysteine and protein persulfidation ultimately affecting mitochondrial respiration.[48] Therefore, we analyzed the sequences of *Cars2*-FL and *Cars2*-AT to identify whether these alternative transcripts have coding potential using the *Coding-Potential Assessment* Tool (CPAT). CPAT predicted both isoforms to be coding (coding potential > 0.99), annotating the open reading frame correlating with the UniProt reference amino acid sequence to *Cars2*-FL and predicting an N terminally truncated protein isoform for *Cars2*-AT using the same open reading frame but missing the first 244 aa (Figure 7F). As expected for a mitochondrial protein, targetP predicted a mitochondrial localization signal at the N terminus of the full length Cars2 protein, which was missing in the truncated protein isoform (Figure 7F). The truncated Cars2 protein isoform was further predicted to lack parts of the conserved binding sites for both $Zn^{2+}$ and pyridoxal phosphate (PLP), indicating a

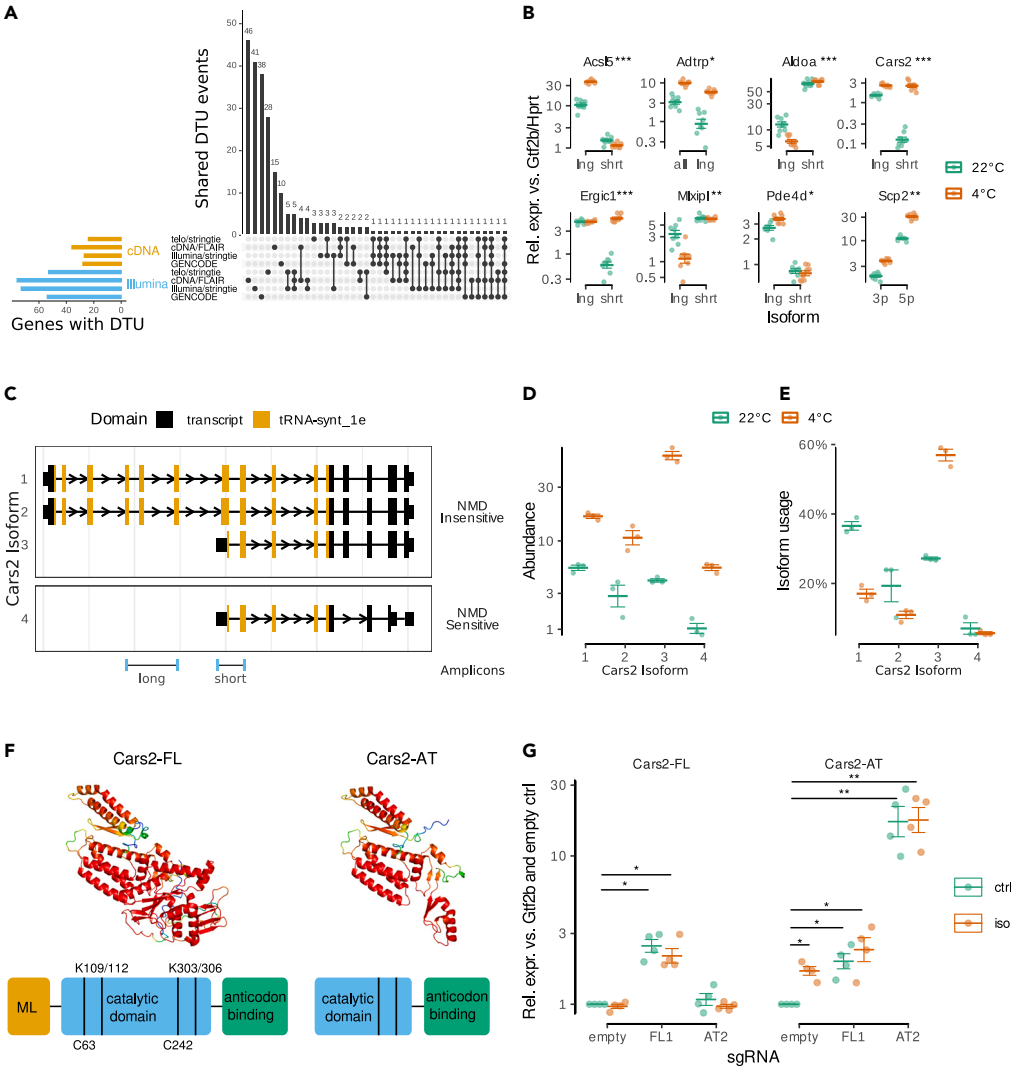



**Figure 7. Differential transcript usage in cold-activated murine iBAT**

(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. Colors represent identified protein domains.

(D and E) Abundance (normalized 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.,[48] and prediction of secondary and tertiary structures of the full length CARS2 and the predicted truncated protein isoform by LocalColabFold. The color 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. Data are represented as mean ± SEM. See also Figures S6–S9.

potential lack of catalytic function (Figure 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 immortalized brown adipocytes (Figure 7G). Using a sgRNA targeting the alternative promoter, it was possible to induce the expression of *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 (Figure 7G). Thus, these results demonstrate that long-read DTU can identify and quantify biologically relevant changes in isoform usage.

## 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,[49–51] 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.[32] 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 outperforms 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.[29,31,32] 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.[29,37] PCR amplification protocols such as TeloPrime typically produce higher sequencing depth than PCR-free methods, increasing coverage which is required for accurate identification of alternative transcripts.[24,29] 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.[29,32]

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; Figure 6A), in line with previous reports.[26] In contrast to StringTie, FLAIR incorporates ChIP/CAGE-Seq data which can be high value because it discriminates between true internal TSS and artifacts 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-treated mice and predicted to encode an N-terminally truncated protein (Figure S6). Cars2 has recently been reported to be involved in sulfur metabolism, which is of importance for mitochondrial morphology and BAT function.[48,52] 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 adipocyte cell model overexpressing *Cars2*-AT and *Cars2*-FL by CRISPR/Cas9 mediated activation of the respective endogenous promoter (Figure 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.[53,54] TeloPrime's strategy for 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 (Figures 7B and 7C): Cellular cAMP levels are also regulated cAMP-specific phosphodiesterases (PDEs) and *Pde4d* regulates lipolysis and thermogenic gene expression.[55] Cold-activated induction of *Pde4d*-long might lead to reduced thermogenic cAMP signaling since *Pde4d*-long controls cAMP levels negatively in a spatial manner due to UCR motifs present in the long *Pde4d* isoform.[56] Another factor regulating the thermogenic β3-AR signaling cascade is *Adtrp*; *Adtrp*-deficient mice are cold-intolerant and have defective thermogenesis.[57] 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.[58] The cold-induced switch in the peroxisome lipid transfer protein *Scp2* in favor of the longer (and enzymatically active) isoform (Figure 7B) may thus support a peroxisome-to-mitochondria lipid signaling hub, supporting mitochondrial uncoupling and thermogenesis.[59] Proper BAT thermogenic function requires cellular protein quality control and removal of misfolded proteins.[60] We here identified DTU in the protein sorting gene *Ergic1*[61] with the shorter isoform being specifically increased upon cold activation (Figures 7B and 7C), 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.[62] Based on our DTU analysis of *Mlxipl*, we speculate that the observed downregulation of the long, intact *Mlxipl*

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.[63] Acsl5 activity increases adiposity, decreases Ucp1 expression and energy expenditure in mice.[63] Here, we show DTU in *Acsl5* giving rise to two transcripts differing in their 5′UTR but otherwise identical protein domain structure (Figures 7B and 7C) with the longer isoform being the almost exclusively expressed variant in cold, suggesting differential control of translation efficiency of this fatty acid channeling 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 (G3P) shown to be increased upon cold exposure in BAT of mice.[64] 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.

### Limitations of the study

There are limitations to this work that should be noted. While our study demonstrates several new isoforms of genes undergoing DTU between control and cold-activated murine BAT, our data are limited to the level of RNA expression. Further studies are needed to investigate whether the newly discovered transcript isoforms presented here affect cold-induced thermogenesis in BAT using gain/loss of function models. Although this study showed an extensive and complex analysis of alternative transcription in cold-activated BAT, other RNA processing events such as alternative 5′capping, or variations in poly-A tail length may also contribute to BAT transcript complexity.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Study animals and housing
  - Brown adipocyte cell culture
- METHOD DETAILS
  - RNA isolation
  - Illumina RNA sequencing
  - ONT library preparation
  - Reverse transcription and qPCR
  - Transcriptome reannotation
  - Annotation of transcripts
  - 7 Cars2 overexpression in cell culture
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Long-read alignment and quantification
  - Short-read alignment and quantification
  - Differential gene and transcript expression analysis
  - Analysis of differential transcript usage

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107190.

## ACKNOWLEDGMENTS

This work was supported by a European Research Council starting grant (TransGen RNA, 675014), a Novo Nordisk Foundation grant (Adiposign, NNF18OC0033444) and a research grant from the Danish Diabetes Academy, which is funded by the Novo Nordisk Foundation (NNF12SA1016522).

ONT sequencing was performed at the Core Facilities of the Medical University of Vienna, a member of VLSI. We thank Markus Jeitler (Core Facilities, Medical University of Vienna) for help with ONT library

preparation and sequencing, and Bjørk Ditlev Marcher Larsen and Ajeetha Josephrajan (Department for Biochemistry and Molecular Biology, University of Southern Denmark) for generating and visualizing the Cars2 structure predictions, respectively. The wt1-SAM brown preadipocyte cell line as well as the empty sgRNA vector have been generated as described in Lundh et al.,[65] and were kind gifts of Brice Emanuelli.

## AUTHOR CONTRIBUTIONS

Conceptualization, C.A.E., M.B., J.-W.K.; Software, C.A.E., S.D., Validation, C.A.E., M.B., J.-W.K.; Formal Analysis, C.A.E.; Investigation, C.A.E., S.K.; Data Curation, C.A.E., S.D.; Visualization, C.A.E.; Writing—Original Draft Preparation, C.A.E.; Writing—Review and Editing, C.A.E., M.B., J.-W.K.; Supervision, M.B., J.-W.K.; Resources, M.B., J.-W.K.; Project Administration, M.B., J.-W.K.; Funding Acquisition, M.B., J.-W.K.

J.-W.K supervised and provided resources for the generation of *in vitro* and *in vivo* samples and the short-read sequencing, M.B. supervised and provided resources for the long-read sequencing. Both contributed equally in the supervision of the data analysis and the preparation of the manuscript.

All authors have read and agreed to the published version of the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interest.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: January 30, 2023
Revised: April 28, 2023
Accepted: June 16, 2023
Published: June 23, 2023

## REFERENCES

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

2. Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413–1415. https://doi.org/10.1038/ng.259.

3. Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476. https://doi.org/10.1038/nature07509.

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

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

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

7. Weatheritt, R.J., Sterne-Weiler, T., and Blencowe, B.J. (2016). The ribosome-engaged landscape of alternative splicing. Nat. Struct. Mol. Biol. *23*, 1117–1123. https://doi.org/10.1038/nsmb.3317.

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

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

10. Robinson, E.K., Jagannatha, P., Covarrubias, S., Cattle, M., Smaliy, V., Safavi, R., Shapleigh, B., Abu-Shumays, R., Jain, M., Cloonan, S.M., et al. (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.

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

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

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

14. Yi, X., Yang, Y., Wu, P., Xu, X., and Li, W. (2020). Alternative splicing events during adipogenesis from hMSCs. J. Cell. Physiol. *235*, 304–316. https://doi.org/10.1002/jcp.28970.

15. Chi, Y.-L., and Lin, J.-C. (2018). RBM4a modulates the impact of PRDM16 on development of brown adipocytes through an alternative splicing mechanism. Biochim. Biophys. Acta Mol. Cell Res. *1865*, 1515–1525. https://doi.org/10.1016/j.bbamcr.2018.08.001.

16. Li, D., Zhang, F., Zhang, X., Xue, C., Namwanje, M., Fan, L., Reilly, M.P., Hu, F., and Qiang, L. (2016). Distinct functions of PPARγ isoforms in regulating adipocyte plasticity. Biochem. Biophys. Res. Commun. *481*, 132–138. https://doi.org/10.1016/j.bbrc.2016.10.152.

17. Byrne, A., Cole, C., Volden, R., and Vollmers, C. (2019). Realizing the potential of full-length

transcriptome sequencing. Philos. Trans. R. Soc. Lond. B Biol. Sci. *374*, 20190097. https://doi.org/10.1098/rstb.2019.0097.

18. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527. https://doi.org/10.1038/nbt.3519.

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

20. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., et al. (2009). Real-Time DNA Sequencing from Single Polymerase Molecules. Science *323*, 133–138. https://doi.org/10.1126/science.1162986.

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

22. Garalde, D.R., Snell, E.A., Jachimowicz, D., Sipos, B., Lloyd, J.H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., et al. (2018). Highly parallel direct RNA sequencing on an array of nanopores. Nat. Methods *15*, 201–206. https://doi.org/10.1038/nmeth.4577.

23. Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. Nat. Rev. Genet. *20*, 631–656. https://doi.org/10.1038/s41576-019-0150-2.

24. Glinos, D.A., Garborcauskas, G., Hoffman, P., Ehsan, N., Jiang, L., Gokden, A., Dai, X., Aguet, F., Brown, K.L., Garimella, K., et al. (2022). Transcriptome variation in human tissues revealed by long-read sequencing. Nature *608*, 353–359. https://doi.org/10.1038/s41586-022-05035-y.

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

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

27. Tang, A.D., Soulette, C.M., van Baren, M.J., Hart, K., Hrabeta-Robinson, E., Wu, C.J., and Brooks, A.N. (2020). Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. Nat. Commun. *11*, 1438. https://doi.org/10.1038/s41467-020-15171-6.

28. Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I., et al. (2021). GENCODE 2021. Nucleic Acids Res. *49*, D916–D923. https://doi.org/10.1093/nar/gkaa1087.

29. Chen, Y., Davidson, N.M., Wan, Y.K., Patel, H., Yao, F., Low, H.M., Hendra, C., Watten, L., Sim, A., Sawyer, C., et al. (2021). A systematic benchmark of Nanopore long read RNA sequencing for transcript level analysis in human cell lines. https://doi.org/10.1101/2021.04.21.440736.

30. Ip, C.L.C., Loose, M., Tyson, J.R., de Cesare, M., Brown, B.L., Jain, M., Leggett, R.M., Eccles, D.A., Zalunin, V., Urban, J.M., et al. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. https://doi.org/10.12688/f1000research.7201.1.

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

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

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

34. Klever, M.-K., Sträng, E., Jungnitsch, J., Melo, U.S., Hetzel, S., Dolnik, A., Schöpflin, R., Schrezenmeier, J.F., Blau, O., Westermann, J., et al. (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.

35. 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., et al. (2021). Full-length transcript sequencing of human and mouse cerebral cortex identifies widespread isoform diversity and alternative splicing. Cell Rep. *37*, 110022. https://doi.org/10.1016/j.celrep.2021.110022.

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

37. Soneson, C., Yao, Y., Bratus-Neuenschwander, A., Patrignani, A., Robinson, M.D., and Hussain, S. (2019). A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes. Nat. Commun. *10*, 3359–3414. https://doi.org/10.1038/s41467-019-11272-z.

38. 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., et al. (2022). Long read sequencing reveals novel isoforms and insights into splicing regulation during cell state changes. BMC Genom. *23*, 42. https://doi.org/10.1186/s12864-021-08261-2.

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

40. Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. *4*, 1521. https://doi.org/10.12688/f1000research.7563.2.

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

42. Shumate, A., Wong, B., Pertea, G., and Pertea, M. (2022). Improved transcriptome assembly using a hybrid of long and short reads with StringTie. PLoS Comput. Biol. *18*, e1009730. https://doi.org/10.1371/journal.pcbi.1009730.

43. De Paoli-Iseppi, R., Gleeson, J., and Clark, M.B. (2021). Isoform Age - Splice Isoform Profiling Using Long-Read Technologies. Front. Mol. Biosci. *8*, 711733. https://doi.org/10.3389/fmolb.2021.711733.

44. Ringeling, F.R., Chakraborty, S., Vissers, C., Reiman, D., Patel, A.M., Lee, K.-H., Hong, A., Park, C.-W., Reska, T., Gagneur, J., et al. (2022). Partitioning RNAs by length improves transcriptome reconstruction from short-read RNA-seq data. Nat. Biotechnol. *40*, 741–750. https://doi.org/10.1038/s41587-021-01136-7.

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

46. Engelhard, C.A., Huang, C., Khani, S., Kasparek, P., Prochazka, J., Rozman, J., Reguera, D.P., Sedlacek, R., and Kornfeld, J.-W. (2022). Comprehensive Transcriptional Profiling and Mouse Phenotyping Reveals Dispensable Role for Adipose Tissue Selective Long Noncoding RNA Gm15551. Noncoding. RNA *8*, 32. https://doi.org/10.3390/ncrna8030032.

47. Rajendran, V., Kalita, P., Shukla, H., Kumar, A., and Tripathi, T. (2018). Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. Int. J. Biol. Macromol. *111*, 400–414. https://doi.org/10.1016/j.ijbiomac.2017.12.157.

48. Akaike, T., Ida, T., Wei, F.-Y., Nishida, M., Kumagai, Y., Alam, M.M., Ihara, H., Sawa, T., Matsunaga, T., Kasamatsu, S., et al. (2017). Cysteinyl-tRNA synthetase governs cysteine polysulfidation and mitochondrial bioenergetics. Nat. Commun. *8*, 1177. https://doi.org/10.1038/s41467-017-01311-y.

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

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

51. Wang, W., and Seale, P. (2016). Control of brown and beige fat development. Nat. Rev. Mol. Cell Biol. *17*, 691–702. https://doi.org/10.1038/nrm.2016.96.

52. 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., and Branco, L.G.S. (2018). Endogenous peripheral hydrogen sulfide is propyretic: its permissive role in brown adipose tissue thermogenesis in rats. Exp. Physiol. *103*, 397–407. https://doi.org/10.1113/EP086775.

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

54. Tomita, H., Shakkottai, V.G., Gutman, G.A., Sun, G., Bunney, W.E., Cahalan, M.D., Chandy, K.G., and Gargus, J.J. (2003). Novel truncated isoform of SK3 potassium channel is a potent dominant-negative regulator of SK currents: implications in schizophrenia. Mol. Psychiatr. *8*, 524–535. https://doi.org/10.1038/sj.mp.4001271.

55. Vezzosi, D., and Bertherat, J. (2011). Phosphodiesterases in endocrine physiology and disease. Eur. J. Endocrinol. *165*, 177–188. https://doi.org/10.1530/EJE-10-1123.

56. Lynch, M.J., Baillie, G.S., and 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. Biochem. Soc. Trans. *35*, 938–941. https://doi.org/10.1042/BST0350938.

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

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

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

60. 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., et al. (2018). Brown adipose tissue thermogenic adaptation requires Nrf1-mediated proteasomal activity. Nat. Med. *24*, 292–303. https://doi.org/10.1038/nm.4481.

61. Joshi, A.S., Zhang, H., and Prinz, W.A. (2017). Organelle biogenesis in the endoplasmic reticulum. Nat. Cell Biol. *19*, 876–882. https://doi.org/10.1038/ncb3579.

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

63. 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., and Greenberg, A.S. (2016). Acyl CoA synthetase 5 (ACSL5) ablation in mice increases energy expenditure and insulin sensitivity and delays fat absorption. Mol. Metabol. *5*, 210–220. https://doi.org/10.1016/j.molmet.2016.01.001.

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

65. Lundh, M., Pluciñska, K., Isidor, M.S., Petersen, P.S.S., and Emanuelli, B. (2017). Bidirectional manipulation of gene expression in adipocytes using CRISPRa and siRNA. Mol. Metabol. *6*, 1313–1320. https://doi.org/10.1016/j.molmet.2017.07.001.

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

67. Köster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. Bioinformatics *28*, 2520–2522. https://doi.org/10.1093/bioinformatics/bts480.

68. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. j. *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

69. Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics *32*, 3047–3048. https://doi.org/10.1093/bioinformatics/btw354.

70. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

71. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

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

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

74. Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763. https://doi.org/10.1093/bioinformatics/14.9.755.

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

76. Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. https://doi.org/10.12688/f1000research.23297.1.

77. Tardaguila, M., de la Fuente, L., Marti, C., Pereira, C., Pardo-Palacios, F.J., Del Risco, H., Verheggen, K., Ferrell, M., Mellado, M., et al. (2018). SQANTI: extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. Genome Res. *28*, 396–411. https://doi.org/10.1101/gr.222976.117.

78. Jain, C., Dilthey, A., Koren, S., Aluru, S., and Phillippy, A.M. (2017). A Fast Approximate Algorithm for Mapping Long Reads to Large Reference Databases. In Research in Computational Molecular Biology Lecture Notes in Computer Science, S.C. Sahinalp, ed. (Springer International Publishing), pp. 66–81. https://doi.org/10.1007/978-3-319-56970-3_5.

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

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

81. Nowicka, M., and Robinson, M.D. (2016). DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. F1000Res. *5*, 1356. https://doi.org/10.12688/f1000research.8900.2.

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

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

84. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat. Methods *19*, 679–682. https://doi.org/10.1038/s41592-022-01488-1.

85. 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., et al. (2019). Nanopore native RNA sequencing of a human poly(A) transcriptome. Nat. Methods *16*, 1297–1305. https://doi.org/10.1038/s41592-019-0617-2.

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

87. 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., et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. *44*, D279–D285. https://doi.org/10.1093/nar/gkv1344.

88. 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., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588. https://doi.org/10.1038/nature14136.

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

90. Zhu, A., Ibrahim, J.G., and Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics *35*, 2084–2092. https://doi.org/10.1093/bioinformatics/bty895.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Biological samples** | | |
| Murine C57BL/6N interscapular brown adipose tissue from 20 week old, male mice, cold treated for 24 h or control | This paper | |
| **Chemicals, peptides, and recombinant proteins** | | |
| EDTA | Sigma | Cat #: E5134 |
| RPMI 1640 Medium, GlutaMAX™ Supplement | Gibco/ThermoFisher Scientific | Cat #: 61,870,036 |
| PBS, pH 7.4 | Gibco/ThermoFisher Scientific | Cat #: 61,870,036 |
| Fetal Bovine Serum | Sigma | Cat #: F7524 |
| Indomethacin | Sigma | Cat #: I7378 |
| Rosiglitazone | Sigma | Cat #: R2408 |
| Dexamethasone | Sigma | Cat #: D8893 |
| Insulin | Sigma | Cat #: I9278 |
| 3-Isobutyl-1-methylxanthin (IBMX) | Sigma | Cat #: I5879 |
| Penicillin-streptomycin | Lonza | Cat #: DE17/602E |
| Opti-MEM II | ThermoFisher Scientific | Cat #: 31985047 |
| DMEM | ThermoFisher Scientific | Cat #: 11960085 |
| Glutamin | ThermoFisher Scientific | Cat #: 25030024 |
| TransIT-X2 | Mirus | Cat #: MIR 6005 |
| Trypsin | ThermoFisher Scientific | Cat #: 25200-056 |
| Gelatine | Sigma | Cat #: G1393 |
| Sodium Acetate | Sigma | Cat #: S2889 |
| Glycogen | Millipore | Cat #: 361507 |
| **Critical commercial assays** | | |
| High Capacity cDNA Reverse Transcription Kit | Applied Biosystems | Cat #:4368814 |
| Flow Cell (R9.4.1) | Oxford Nanopore Technologies | Cat #: FLO-MIN106 |
| Direct RNA Sequencing Kit | Oxford Nanopore Technologies | Cat #: SQK-RNA002 |
| Direct cDNA Sequencing Kit | Oxford Nanopore Technologies | Cat #: SQK-DCS109 |
| Barcoding kit | Oxford Nanopore Technologies | Cat #: EXP-NBD104 |
| TeloPrime Full-Length cDNA Amplification Kit | Lexogen | Cat #: 0.13.08 |
| The NEBNext Ultra II Directional RNA Library Prep Kit | New England Biolabs | Cat #: E7760S |
| NovaSeq 6000 S1 Reagent Kit v1.5 (100 cycles) | Illumina | Cat #: 20028319 |
| GenElute mRNA Miniprep Kit | Sigma | Cat #: MRN 10 |
| **Deposited data** | | |
| Direct cDNA sequencing data | This paper | GEO: GSE212572 |
| TeloPrime cDNA sequencing data | This paper | GEO: GSE212571 |
| Direct RNA sequencing data | This paper | GEO: GSE212570 |
| Raw Illumina sequencing data | This paper | GEO: GSE212569 |
| murine iBAT H3K4me3 ChIP-Seq data | Engelhard et al.[46] | GEO: GSE200651 |
| FANTOM5 CAGE peaks | Abugessaisa et al.[66] | mm10.cage_peak_phase1and2combined_coord.bed |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Mouse reference transcript sequences | Gencode | ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M22/gencode.vM22.transcripts.fa.gz |
| Mouse reference annotation | Gencode | ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M22/gencode.vM22.primary_assembly.annotation.gtf.gz |
| Mouse reference genome (GRCm38.p6) | Gencode | ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M22/GRCm38.primary_assembly.genome.fa.gz |
| Code to reproduce analysis and figures | This paper | Zenodo: https://doi.org/10.5281/zenodo.7990800 |
| **Experimental models: Cell lines** | | |
| wt1-SAM brown preadipocytes | Prof. Brice Emanuelli | |
| **Experimental models: Organisms/strains** | | |
| Mouse: C57BL/6N, Wild Type | Charles River | |
| **Recombinant DNA** | | |
| sgRNA(MS2) cloning backbone | Addgene | Cat #: 61424 |
| **Software and algorithms** | | |
| fastqc v0.11.9 | Andrews, 2010 | https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ |
| Snakemake | Köster and Rahmann,[67] | https://snakemake.readthedocs.io/en/stable/ |
| cutadapt v3.7 | Martin[68] | https://cutadapt.readthedocs.io/en/stable/index.html |
| multiqc v1.12 | Ewels et al.[69] | https://multiqc.info/ |
| samtools v1.12 | Li et al.[70] | https://www.htslib.org/ |
| bedtools v2.30 | Quinlan and Hall,[71] | https://bedtools.readthedocs.io/en/latest/index.html |
| nanofilt v2.3.0 | De Coster et al.[72] | https://github.com/wdecoster/nanofilt |
| minimap2 v2.24 | Li et al.[73] | https://github.com/lh3/minimap2#cs |
| salmon v1.8.0 | Patro et al.[19] | https://combine-lab.github.io/salmon/ |
| pfam_scan v1.6 | Aziele | http://xfam.org/ |
| hmmer v3.3.2 | Eddy et al.[74] | http://hmmer.org/ |
| STAR v2.7.8a | Dobin et al.[75] | https://github.com/alexdobin/STAR |
| FLAIR v1.5 | Tang et al.[27] | https://flair.readthedocs.io/en/latest/ |
| stringtie v2.2.0 | Shumate et al.[42] | https://ccb.jhu.edu/software/stringtie/ |
| gffcompare v0.12.6 | Pertea and Pertea,[76] | https://ccb.jhu.edu/software/stringtie/gffcompare.shtml |
| gffread v0.12.7 | Pertea and Pertea,[76] | http://ccb.jhu.edu/software/stringtie/gff.shtml |
| sqanti2 v7.4.0 | Tardaguila et al.[77] | https://github.com/Magdoll/SQANTI2 |
| mashmap v2.0 | Jain et al.[78] | https://github.com/marbl/MashMap |
| R v4.1.2 | R Core Team, 2021 | https://cran.r-project.org/ |
| tximport v 1.22.0 | Soneson et al.[40] | https://bioconductor.org/packages/release/bioc/html/tximport.html |
| GenomicAlignments v1.30.0 | Lawrence et al.[79] | https://bioconductor.org/packages/release/bioc/html/GenomicAlignments.html |
| DESeq2 v1.34.0 | Love et al.[80] | https://bioconductor.org/packages/release/bioc/html/DESeq2.html |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| DRIMSeq v1.22.0 | Nowicka and Robinson,[81] | https://bioconductor.org/packages/release/bioc/html/DRIMSeq.html |
| IsoformSwitchAnalyzeR v1.16.0 | Vitting-Seerup and Sandelin,[82] | https://bioconductor.org/packages/release/bioc/html/IsoformSwitchAnalyzeR.html |
| CRISPick | Doench et al.[83] | https://portals.broadinstitute.org/gppx/crispick/public |
| LocalColabFold | Mirdita et al.[84] | https://colabfold.mmseqs.com/ |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Prof. Dr. Jan-Wilhelm Kornfeld (janwilhelmkornfeld@bmb.sdu.dk).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- RNA-seq data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Any other type of data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Study animals and housing

For maintenance, 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. For the experiment, 20 weeks old male wild type C57BL/6N mice were singly housed at 4°C (cold treatment) or at 22°C to 24°C (control group) for a period of 24 h prior to harvesting adipose tissues.

All animal experiments were approved by the Danish Miljø- og Fødevarestyrelsen (license 2018-15-201-01548) and conformed to the relevant regulatory standards.

### Brown adipocyte 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 confluence, 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.

## METHOD DETAILS

### RNA isolation

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.[85]

### Illumina RNA sequencing

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′-TCTCAGGCGTTT TTTTTTTTTTTTTT-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

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 primer sets were designed using Primer3Plus. They either include one primer overlapping an exon-exon junction, or the two primers are placed in different exons to exclude the amplification of genomic DNA. Primer sets were designed against sets of isoforms as indicated in Figures S7–S9 (amplicon tracks). Sequences of the primers and chromosomal start/end positions of the amplicons are listed in Table S1.

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 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 Cq values calculated as the second derivative maximum. Genes of interest were normalised against housekeeper genes using the ΔCq method.

### Transcriptome reannotation

For transcriptome reannotation using StringTie42 (v.2.2.0), bam files from aligning the reads to the genome as described above were provided to stringtie. Additionally, the reference annotation in gtf format

(-G, GENCODE M22) was provided to stringtie and a splice junction cutoff (-j) of 10, a minimum coverage (-c) of 1 and a minimum transcript fraction (-f) of 1 % were used. For ONT long-read runs, the –mix mode was used, additionally providing StringTie with the short-read bam files of the same sample. Afterwards, all re-annotations 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). Reference annotation was not provided to StringTie in the merge step.

For the reannotation using FLAIR27 (v1.5), bam files from alignment against the genome as described above were first converted to bed format using bam2Bed12.py from the FLAIR suite. Splice junction coverage files from aligning the short-read data to the genome using STAR (SJ.out.tab) were filtered for a minimum splice junction coverage of 10 using a custom R script. For each sample, the bed files and the corresponding filtered short-read based junction data was provided to flair.py correct, additionally providing the reference genome in fasta format (-g) and a table of chromosome sizes (-c, created from the genome fasta using samtools faidx). Subsequently, the corrected reads in psl format were concatenated into a single psl file using gnu cat and collapsed into a transcriptome reannotation by FLAIR collapse using a minimum coverage (-s) of 3 and setting the –stringent flag, additionally providing the genome in fasta format (-g), the reference annotation in gtf format (-f) and the reads from all samples of the respective library prep method in fastq format (-t). In order to mark true transcriptional start sites, a combined bed file from FANTOM5 CAGE peaks67 and iBAT H3K4me3 peaks48 was provided. For the TeloPrime data, the –trust_ends flag was additionally set.

The reannotated transcripts were compared to the reference annotation using SQANTI2[77] (v7.4.0). The overlap between annotated known (classcode =, m,and c) and novel transcripts in the different datasets and the reference annotation was calculated using the tracking file from gffcompare.

### Annotation of transcripts

Annotation of transcript isoforms including open reading frames, nonsense mediated decay,[86] functional protein domains (pfam[87]) and splice type analysis was done using IsoformSwitchAnalyzeR[82] (v1.16.0).

Cars2 protein fold predictions were generated in LocalColabFold using standard parameters72. Computation of the models was performed on the UCloud interactive HPC system, which is managed by the eScience Center at the University of Southern Denmark.

### 7 Cars2 overexpression in cell culture

For *in vitro* gain of function studies using the wt1-SAM cell line, single guide RNAs (sgRNAs) were designed using CRISPick73 and cloned into the sgRNA(MS2) cloning backbone (addgene 61424) as described by Konermann.[88] 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, 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.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Long-read alignment and quantification

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 minimap275 (v2.24). Reads were subsequently filtered for an average PHRED score > 7 using NanoFilt76 (v2.3.0). GenomicAlignments77 (1.30.0) 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 Soneson37.

### Short-read alignment and quantification

Illumina short-reads were quality filtered using cutadapt78 (v.3.7; -q 28 -m 30) and mapped to the genome (GRCm38.p6) using STAR79 (v2.7.8a). Transcript level quantification was done using salmon19 in selective alignment mode (GENCODE M22; v1.8.0). Gene level abundance estimation was done using tximport40 (v1.22.0).

### Differential gene and transcript expression analysis

Differential gene and transcript expression analysis was done using DESeq2 with $H_0$: log2FC > 0.580 (v1.34.0). Adjustment for multiple testing was done using the *s*-value method proposed by Stephens[89] implemented in apeglm[90] and *s*-values < 0.05 were considered significant.

### 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 DRIMSeq83 (v.1.22.0). 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.