

At the Dawn of the Transcriptomic Medicine

Journal:	<i>Experimental Biology and Medicine</i>
Manuscript ID	EBM-20-AMI-0600.R1
Manuscript Type:	Annual Minireview Issue 2021
Date Submitted by the Author:	12-Aug-2020
Complete List of Authors:	Koks, Gea; Prion Ltd Pfaff, Abigail; Murdoch University Bubb, Vivien; University of Liverpool Quinn, John; University of Liverpool Koks, Sulev; Perron Institute for Neurological and Translational Science, ; Murdoch University,
Keywords:	TRANSCRIPTOME, RNA-Seq, Gene Expression Profiling, Precision Medicine, GENOMICS, Molecular Targeted Therapy
Abstract:	<p>Progress in genomic analytical technologies has improved our possibilities to obtain information regarding DNA, RNA and their dynamic changes that occur over time or in response to specific challenges. This information describes the blueprint for cells, tissues and organisms and has fundamental importance for all living organisms. This review focuses on the technological challenges to analyse the transcriptome and what is the impact of transcriptomics on precision medicine. The transcriptome is a term that covers all RNA present in cells and a substantial part of it will never be translated into protein but is nevertheless functional in determining cell phenotype. Recent developments in transcriptomics have challenged the fundamentals of the central dogma of biology by providing evidence of pervasive transcription of the genome. Such massive transcriptional activity is challenging the definition of a gene and especially the term "pseudogene" that has now been demonstrated in many examples to be both transcribed and translated. We also review the common sources of biomaterials for transcriptomics and justify the suitability of whole blood RNA as the current optimal analyte for clinical transcriptomics. At the end of the review, a brief overview of the clinical implications of transcriptomics in clinical trial design and clinical diagnosis is given. Finally, we introduce the transcriptome as a target for modern drug development as a tool for extending our capacity for precision medicine in multiple diseases.</p>

SCHOLARONE™
Manuscripts

At the Dawn of the Transcriptomic Medicine

Gea Koks¹, Abigail L. Pfaff^{2,3}, Vivien J. Bubb⁴, John P. Quinn⁴, Sulev Koks^{2,3}

¹ Prion Ltd, Tartu 50410, Estonia;

² Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Murdoch
6150, Perth, WA, Australia;

³ The Perron Institute for Neurological and Translational Science, Nedlands 6009, Perth, WA,
Australia;

⁴ Department of Pharmacology and Therapeutics, University of Liverpool, Liverpool L69 3BX,
UK

Address for Correspondence:

Sulev Koks

8 Verdun Street, Nedlands 6009, WA, Australia

Phone: +61 (0) 8 6457 0313

e-mail: sulev.koks@perron.uwa.edu.au

Running Title

Transcriptomic Medicine

Abstract

Progress in genomic analytical technologies has improved our possibilities to obtain information regarding DNA, RNA and their dynamic changes that occur over time or in response to specific challenges. This information describes the blueprint for cells, tissues and organisms and has fundamental importance for all living organisms. This review focuses on the technological challenges to analyse the transcriptome and what is the impact of transcriptomics on precision medicine. The transcriptome is a term that covers all RNA present in cells and a substantial part of it will never be translated into protein but is nevertheless functional in determining cell phenotype. Recent developments in transcriptomics have challenged the fundamentals of the central dogma of biology by providing evidence of pervasive transcription of the genome. Such massive transcriptional activity is challenging the definition of a gene and especially the term “pseudogene” that has now been demonstrated in many examples to be both transcribed and translated. We also review the common sources of biomaterials for transcriptomics and justify the suitability of whole blood RNA as the current optimal analyte for clinical transcriptomics. At the end of the review, a brief overview of the clinical implications of transcriptomics in clinical trial design and clinical diagnosis is given. Finally, we introduce the transcriptome as a target for modern drug development as a tool for extending our capacity for precision medicine in multiple diseases.

Keywords

Transcriptome, RNA-Seq, Gene Expression Profiling, Precision Medicine, Genomics, Molecular Targeted Therapy

Impact statement

This review describes the impact of transcriptomics on experimental biology and its integration into medical practice. Transcriptomics is an essential part of modern biomedical research based on highly sophisticated and reliable technology. Transcriptomics can aid clinical practice and improve the precision of clinical diagnoses and decision-making by complementing existing clinical best practice. The power of which will be increased when combined with genomic variation from genome wide association studies and next generation sequencing. We are witnessing the implementation of RNA-based technologies in clinical practice that will eventually lead to the establishment of transcriptional medicine as a routine tool in diagnosis.

Introduction

Since the identification of the structure of proteins and nucleic acids and the mechanisms of gene expression, the central concept of biology has underpinned our understanding of gene function¹. According to this concept, the information in the cell is from DNA to RNA and subsequently translated into proteins. Therefore, the function of genes should be analysed only by their ability to produce proteins and that proteins define phenotype. One field of research, transcriptomics, has revolutionised this central biological concept. Discovery of the abundance and complexity of RNA dynamics and function dramatically changed our understanding about the role of RNA, apart from encoding proteins, and challenged gene-centric approach to explain the function of genome². Transcriptome is a collective term describing all RNAs produced by a single cell, by a population of cells or tissue³. Recent progress in analytical technologies has unveiled the complexity of the regulation of the transcriptome. The transcriptome is the primary product of the genome and therefore analysis of the transcriptome provides primary information for functional genomics.

The human genome

One of original and the most remarkable results of the human genome project was the discovery that only 1.2% of the human genome encodes proteins and was therefore considered as functional and meaningful⁴. This is also reflected in the early search for genetic variation associated with a specific disease focusing on DNA sequences solely in exons. The number of genes, protein-coding elements, was discovered to be around 30,000, a much smaller number than predicted and similar to that found in several other species⁴. Based on these findings the rest of the genome was initially termed as junk DNA. However, additional studies have identified that most of the DNA has function, not only for genome

1
2
3 structure and packaging, but also to form the complexity of the molecular networks
4
5 underpinning the diversity of cell function. Early studies, after the identification of individual
6
7 chromosome sequences, indicated that genomic sequences were transcribed at least as
8
9 much as an order of magnitude more than accounted for by the predicted gene models².
10
11 Similarly, the term “pseudogene” that implies that is not a real gene and considered as a
12
13 remnant of evolution or “genomic fossil”⁵. It is now demonstrated that most of the
14
15 pseudogenes are transcribed and translated into proteins challenging that definition of
16
17 “pseudogene”⁶. Cap-analysis gene expression (CAGE) technology enabled the identification
18
19 of at least 180,000 transcripts in the mammalian genome and it appeared that the majority
20
21 of the genome is transcribed ⁷. At least 60% of the genome has been described as a
22
23 transcriptional forest, where transcription is performed from both strands of the same DNA
24
25 region without gaps ⁷. The most remarkable project in this field is known as an Encyclopedia
26
27 of DNA Elements or ENCODE for short. Based on ENCODE findings at least 80% of genome is
28
29 actively transcribed and this number is considered to be conservative⁸. Such data requires
30
31 we review our interpretation of genome function and regulation and how that is utilised in
32
33 clinical translation.
34
35
36
37
38
39
40
41
42
43
44

45 Transcriptome, transcriptomics and transcriptome profiling

46 Transcriptome is a collection of the RNAs (transcripts) that single cell or tissue can produce,
47
48 and it contains all types of RNAs⁹. Transcriptomics is the study of the transcriptome;
49
50 analysing RNA and its different subcategories (mRNA, micro-RNA, non-coding RNA, etc) to
51
52 identify changes in expression and its functional impact. Although transcriptomics focusses
53
54 on content and transcript expression levels, it also includes the analysis of transcriptional
55
56 regulation. The transcriptome can be studied by different methods, however the most
57
58
59
60

1
2
3 common options are genechips (to measure gene expression on microarray platform) and
4
5 RNA sequencing (RNA-seq)⁹. Gene expression arrays initially focused solely on polyA
6
7 purified RNA that encode proteins. Moreover, genechips also suffer from the requirement
8
9 to be pre-designed, i.e. the content on the array is based on our pre-existing knowledge of
10
11 predominantly exons that can be easily identified in genome sequence data^{10, 11}. Therefore,
12
13 genechips give us a snapshot of the transcriptional changes of mRNA, but this snapshot is
14
15 rather limited. More recent arrays (transcript based and tiling arrays) can give very
16
17 comprehensive information about the transcriptional changes, nevertheless the genechips
18
19 are inherently bound to pre-existing knowledge and do not provide information about the
20
21 sequences of the transcripts^{12,10}. Only a few genechip versions are capable of identifying
22
23 alternative splicing and specialised chip design is required to analyse such as micro-RNAs¹².
24
25 But the sequence information is lost in results files, and this is where the RNA-sequencing
26
27 has clear advantage allowing for more detailed analysis to detect alternative splicing, intron
28
29 retention and other events reflecting alterations in transcriptome regulation and the other
30
31 classes of RNA. Therefore, RNA-sequencing has become the main technology for
32
33 transcriptome analysis^{9, 13}.
34
35
36
37
38
39
40
41
42
43
44

45 Sources of the transcriptome

46 Gene expression is both tissue specific and stimulus inducible therefore a key question for
47
48 transcriptome analysis is the source of the tissue or cell type for analysis. The most common
49
50 and easiest to justify is the primary tissue that is affected by pathological processes. This is
51
52 based on the assumption that we know what tissue is affected and we have some
53
54 preliminary understanding what the timeline and mechanisms of the pathological changes
55
56 are. However, this assumption can be deceiving. For example, with CNS disorders, it is
57
58
59
60

1
2
3 difficult to determine which region or cell type is involved and also whether the pathological
4
5 hallmarks of the disease were initiated by dysfunction in another brain regions, or
6
7 periphery, many years before. As brain tissue is only accessible as post-mortem tissue, the
8
9 changes in the transcriptome could arise from selective alteration of gene expression by the
10
11 post-mortem time rather than in response to living with a chronic age-dependent disease
12
13 occurring over a long time period ^{14, 15, 16}. In case of neurodegenerative diseases, this may
14
15 mean that we miss the molecular pathological changes that initiate the degenerative
16
17 process. The same is similar for other chronic age-dependent disease such as arthritis or
18
19 heart disease. The cells that are targeted by primary pathology are often dead or have a
20
21 significantly altered phenotype from those that represent the key pathological transitions.
22
23 Some of the problems of addressing transcriptomics in the CNS are outlined below.
24
25 Firstly, recognised issues with the use of biobanked tissue samples that would affect
26
27 transcriptomics include the heterogeneity of the samples, reliability of the diagnoses and
28
29 variability in the quality control measures ¹⁷. The most drastic example to illustrate
30
31 reliability challenges comes from the biobank having 12,000 samples available for research
32
33 and only 18 of them with the suitable information and quality by the end ¹⁶. While the
34
35 analysis of post-mortem brain samples is still valid and informative from a research point of
36
37 view, the impact of these studies to improve our understanding about neurodegenerative
38
39 disease needs addressed in a broader context ¹⁶. It is difficult to infer causative changes
40
41 from the single time point that is based on the analysis of the tissues where the pathogenic
42
43 processes are completed.
44
45

46
47 Secondly, subjects may have used drugs for a long time and depending on the course of the
48
49 disease the treatment schedules can be quite different between patients ¹⁷. Moreover, it is
50
51 quite realistic to assume that the subjects have had comorbidities and taken drugs for those
52
53
54
55
56
57
58
59
60

1
2
3 symptoms as well. Drugs for heart disease and hypertension and statins are quite common
4
5 in the aged population and therefore analysis of the post-mortem samples should most
6
7 certainly take into account the drug history and comorbidities as confounders. This is
8
9 something we do not see very often in studies using post-mortem tissue samples.
10
11

12
13 Thirdly, we need to consider what regions of the tissue is to be analysed. Again, in the case
14
15 of the brain, regional changes in gene expression can be enormous¹⁷. It is a complex tissue
16
17 and choosing the right regions for comparison is often the most important decision for the
18
19 analysis. For example, in the case of targeted mutation mouse models generated by
20
21 homologous recombination the changes in the transcriptome of the brain are regionally
22
23 very different¹⁸⁻²⁰. Targeted mutant mouse lines allow exclusion all confounding factors and
24
25 careful matching of the study subjects for the genetically engineered mutations. However,
26
27 even after the perfect matching for confounders, the deletion of the single gene induced
28
29 enormously different changes in transcriptome in the different regions of the brain²⁰. Only
30
31 the lack of the expression of the deleted gene was the similar result between the different
32
33 brain regions²⁰. In addition to the regional difference in the brain tissue, genomic locus of
34
35 the gene has also to be considered. We have analysed the transcriptome of the Wolfram
36
37 syndrome mutant mice with the deletion of the *Wfs1* gene and identified significant
38
39 confounding effect from the genomic locus of the targeted gene¹⁸. This locus-specific or
40
41 genomic context effect means that even a single gene targeting, or deletion can induce the
42
43 complex changes in the transcriptome that are not caused by the function of the gene, but
44
45 by its location. Mouse models enable controlling for gender, age and environmental
46
47 differences, providing the ideal study design conditions, but cannot avoid genomic
48
49 background effect, "congenic footprint"²¹. This effect needs to be taken into account and
50
51 with appropriate adjustment the functionally meaningful differences can be identified²². All
52
53
54
55
56
57
58
59
60

1
2
3 this illustrates how diverse the transcriptome is in different brain regions and therefore it is
4
5 challenging to design studies with multiple brain regions involved as it is not trivial to
6
7 differentiate between the normal regional and pathologically relevant differences. In
8
9 summary, by analysing post-mortem brains we struggle to obtain the relevant information
10
11 about the mechanisms of the disease and this information does not always help us to design
12
13 better diagnostic tools or drugs.
14
15

16
17 However, analysis of the diseased tissues is important when it is possible during the
18
19 pathogenesis of the disease. Repeated sampling during the course of the disease allows us
20
21 to use the time-dependent causative interaction models. Longitudinal studies are therefore
22
23 the best way to follow disease progression but severely limit the choice of tissue or
24
25 component that can be measured to such as blood, skin, urine and microbiome. This also
26
27 enables the monitoring of changes in the transcriptome during treatment and to compare
28
29 different therapeutic options²³. In more limited cases, surgical removal of tissue during
30
31 medical procedures is another option to access samples for transcriptomic analysis. The
32
33 latter option is the most common for oncological samples and is potentially applicable for
34
35 any surgically treated conditions. If we plan to perform longitudinal transcriptome analysis
36
37 with samples from different time-points, then almost the only viable option is blood
38
39 sampling. Skin sampling can also be alternative for some cases and diagnoses. We have
40
41 shown that skin and blood are useful alternatives even for neurodegenerative diseases like
42
43 Parkinson's disease²⁴⁻²⁶. Both blood and skin showed clear transcriptome differences in the
44
45 case-control design and these tissues could be used for the diagnosis or monitoring the
46
47 progression of the disease. Similarly, urine can be used as a source for transcriptome
48
49 analysis^{27, 28}. However, as usually the cellular content in urine is low, the RNA level is also
50
51
52
53
54
55
56
57
58
59
60

1
2
3 low and that reduces potential of urine or other body fluids as a source for transcriptomics
4
5 28.
6
7
8
9

10 Whole blood versus PBMC transcriptome

11 Blood is a useful and easy to access surrogate tissue for transcriptome analysis, but the use
12
13 of blood requires a few basic decisions. For example, it is possible to analyse whole blood or
14
15 a particular fraction of blood cells. Peripheral Blood Mononuclear Cell (PBMC) separation
16
17 has been one very popular method to isolate cells from the blood and to prepare them for
18
19 RNA analysis. However, the PBMC fraction contains only lymphocytes and monocytes while
20
21 all granulocytes like basophils, eosinophils and neutrophils are depleted. From all white cell
22
23 count, neutrophils constitute 55 to 75% indicating that using of PBMC for transcriptome
24
25 analysis would not give the full picture ²⁹. Isolation of PBMCs covers only 20 to 50% of the
26
27 cellular heterogeneity of the blood. Moreover, PBMC separation itself is a procedure that
28
29 adds an extra uncontrollable variation to the analysis, and this should be avoided. Several
30
31 studies have shown significant differences between the transcriptome profiles between
32
33 PBMC and whole blood ³⁰. It is reported that over 2,000 genes were differentially expressed
34
35 with more than 2-fold difference between PBMC and whole blood from the same individual
36
37 at same time ³¹. Therefore, for transcriptome analysis the whole blood RNA samples have a
38
39 substantial advantage over PBMC or other fractionation.
40
41
42
43
44
45
46
47
48
49

50 Preanalytical considerations

51 Due to the complexity and the volume of the transcriptomics data preanalytical conditions
52
53 have significant impact on the outcome of the analysis. The inadvertent variations can be
54
55 introduced with the sampling of the tissue, during the storage and transportation or by the
56
57 differences in the extraction methods. In addition, as addressed in previous sections, the
58
59
60

1
2
3 sources for RNA can be variable ranging from blood and other body fluids to the tissue
4
5 biopsies, cellular smears and to single cell sorting. All these different approaches require
6
7 standardised protocols to ensure reproducibility and high quality of the analysis. The testing
8
9 and guidelines how to prepare and purify different clinical samples is vital for the further
10
11 implementation of the transcriptomic analysis in clinical practise. RNA extraction can be
12
13 notoriously complicated with variable options available that all can lead to different results
14
15
16
17 ³². Similarly, storage conditions have been shown to impact the quality of RNA and snap-
18
19 frozen samples detect significantly more genes than FFPE samples ³³. This effect was not
20
21 dependent on the time to fixation. Interestingly, miRNA expression was not affected by the
22
23 fixation method and it was comparable between frozen or FFPE samples ³³. In addition,
24
25 purification of the liquid biopsy samples requires an extra effort and a complex workflow ³⁴.
26
27
28
29 As RNA can be purified from different samples, validation studies are required to develop
30
31 standardised protocols that would enable robust and reproducible analysis of transcriptome
32
33
34
35 for various clinical conditions.

36 37 38 Practical utility of transcriptome analysis

39 The transcriptome is a snapshot of molecular events in the cell reflecting the functional
40
41 activity of the genome at a given moment of time and requires a combination of analytical
42
43 tools to describe these molecular changes. Currently, the majority of genomic tools used in
44
45 clinical genomics only consider targeted DNA sequencing and not the transcriptome.
46
47
48 However, there are several examples of how transcriptomic information improves the
49
50 precision of the genomic analysis.
51
52

53
54 The early studies to analyse transcriptomics used variable differential cloning technologies
55
56 based on cDNA library preparation and comparative analysis ³⁵. One of these methods,
57
58 cDNA Representational Difference Analysis (cDNA-RDA), was used to identify differential
59
60

1
2
3 expression in pancreatic cancer³⁶. cDNA-RDA was proven to be a highly efficient and
4
5 reproducible method that has been used in various models and organisms^{37,38}. While the
6
7 method itself was laborious and difficult to use for larger sample numbers, it clearly had its
8
9 advantage as a hypothesis-free approach to observe transcriptional changes³⁹. As the
10
11 method did not require specific equipment or expensive preparations like gene microchips,
12
13 the method gained popularity and was applied to study variable pathologies or physiological
14
15 responses⁴⁰. At the same time cDNA microarray technology was also developing and
16
17 provided various in-house products These microarrays were based on the cDNA clone
18
19 collections, their amplification and printing (spotting) on to glass slides⁴¹. This technology
20
21 required substantial infrastructure to run and it wasn't widely accessible. Nevertheless, the
22
23 initial studies demonstrated their suitability for pathology and clinical diagnostics in
24
25 particular in cancer where tumour material was available. These studies indicated that breast
26
27 cancers can be classified by their gene expression patterns into subtypes that were not
28
29 identifiable with histological methods alone⁴². The gene expression pattern was not only
30
31 helpful to identify the molecular subtypes of the breast cancers, but also to predict the
32
33 clinical course and outcomes of breast cancer⁴³. These early reports fuelled a myriad of
34
35 similar studies to determine the transcriptional pattern of other tumours to identify
36
37 potential diagnostic or prognostic biomarkers. Gene microarrays became standardised for
38
39 transcriptional studies. The main advantage was the high-throughput analysis of the cDNA
40
41 libraries and as the technology was scalable it was possible to increase sample sizes and the
42
43 power of studies. However, gene arrays still suffered from a biased capture of targets which
44
45 as stated previously were based on exon data or a limited number of non-coding RNAs. The
46
47 latter was partially resolved when Next Generation sequencing (NGS) technologies became
48
49 easily accessible to enable parallel whole genome sequencing (WGS) and RNA-sequencing
50
51
52
53
54
55
56
57
58
59
60

1
2
3 (RNA-seq). Clinical genetics analysis rapidly expanded from exome sequences to a complete
4
5 RNA analysis. RNA-seq is the first technology that enabled complete transcriptome analysis
6
7 covering all different types of RNA subclasses with complete sequence information and
8
9 enables detection of complex profiles from various pathologies⁹.

10
11
12
13 Several examples support the value and the utility of transcriptomics in the complex analysis
14
15 of clinical samples for association to disease. We have analysed the transcriptional profiles
16
17 of osteosarcoma samples from fresh tumours in a paired study design and identified several
18
19 new candidates involved in the development of osteosarcoma²³. Moreover, with similar
20
21 technology we were able to analyse archived formalin fixed paraffin embedded (FFPE)
22
23 samples that gave us the possibility to evaluate the effect of chemotherapy on the
24
25 transcriptional profile. The same dataset provided data regarding repetitive elements that
26
27 were differentially expressed in the malignancy⁴⁴. Repetitive elements can only be
28
29 efficiently analysed using the RNA-seq technology rather than genechips.

30
31
32
33
34
35 Transcriptome analysis can stratify patients who would otherwise be grouped as the same
36
37 disease and this enables biomarker-driven clinical trials to improve their efficacy. Several
38
39 meta-analyses have shown substantial improvement in study outcomes by using the
40
41 biomarker-driven stratification in the study designs⁴⁵. Personalised medicine approaches
42
43 involving biomarkers in study design improved response rates from 5 to 30 percent
44
45 demonstrating the improvement that can be achieved by using a genomics driven
46
47 approach⁴⁵. For example, the Winther trial based on 303 patients utilised genomic-matching
48
49 to personalise their cancer therapy⁴⁶. The study had two arms, one was based only on DNA
50
51 data and the other only on RNA data. This trial introduced several innovative paradigm
52
53 shifts showing an improved therapeutic response with the integration of transcriptomic
54
55 profiling. Most importantly, the transcriptomic arm identified the most suitable solutions for
56
57
58
59
60

1
2
3 the patients with various solid tumours prospectively from the large database of therapies
4
5
6 ⁴⁶. This trial considered patient therapy options at an individual level based on the features
7
8 of person's tumour and not on the results obtained from the aggregation of trials on large
9
10 patient populations. Therapeutic guidance based only on the transcriptomic data resulted in
11
12 the stabilising of disease in 30% of patients ⁴⁶. While not statistically superior from the DNA-
13
14 only approach (26%), it is was a remarkable success considering that the study subjects all
15
16 had advanced cancers with several previous therapies that were unsuccessful.
17
18

19
20 Transcriptomic-guided therapy was considered because the DNA analysis alone does not
21
22 often reveal actionable variants or mutations and RNA analysis could indicate the functional
23
24 consequences. RNA-sequencing served here as an additional analytical tool to describe the
25
26 functional changes in cancer that was in turn used in the therapeutic decision pipeline.
27
28
29 NGS technologies have also changed the ways we analyse Mendelian diseases and made
30
31 whole-exome sequencing (WES) or WGS accessible to identify disease-causing variants.
32
33 However, the success rate for detecting causal changes ranges only from 20 to 30% ⁴⁷. In a
34
35 recent study, the use of RNA-seq analysis yielded diagnostic rate of 35% on previously
36
37 unsolved cases by WGS analysis indicating a marked improvement ⁴⁸. The main advantage of
38
39 RNA-seq is its ability to detect aberrant splicing or disruptive changes in the transcriptional
40
41 regulation that are not detectable with WGS or WES ⁴⁸. This is the evidence to support the
42
43 power of RNA-seq analysis also for Mendelian diseases and shows its clinical applicability in
44
45 this space.
46
47
48
49

50
51 A recent example for the applicability of transcriptome analysis or RNA-based diagnostics
52
53 can be found from the COVID-19 pandemic caused by the RNA-virus SARS-CoV2. The virus is
54
55 only 29,900 bp long and contains 10 genes with gene 5 and 7 being functionally bicistronic
56
57
58 ⁴⁹. Infection is based on the infectious transcriptome and can be viewed as a transcriptome
59
60

1
2
3 infection. Maybe the efficient therapy for viral infections lies in the targeting of the
4
5 transcriptome to affect their transcriptional capacity. Transcriptome based therapies are
6
7 already available for human diseases like Duchenne Muscular Dystrophy or amyloidosis
8
9 showing the potential of the transcriptome based therapeutics⁵⁰⁻⁵². Transcriptome-based
10
11 therapies offer a real systematic opportunity for personalized medicine and it requires
12
13 complex transcriptome analysis as input⁵³. This therapeutic approach can turn the
14
15 information in the transcriptomics into therapeutic options.
16
17
18
19

20 21 Conclusion

22 Transcriptomics is currently a rapidly evolving field with new data to either stand alone or
23
24 integrate with other clinical information to expand and modify the future of health care.
25
26 While current applications are mostly limited to experimental projects, a growing number of
27
28 studies indicate the practical utility of transcriptomics for diagnostics, genomics-driven trial
29
30 design and personalised drug development. Larger clinical validation of such experimental
31
32 hypothesis will allow for accepted clinical usage, indeed blood samples can be taken in
33
34 general practice and sent off for analysis and interpretation centrally before transmission to
35
36 the clinician. Transcriptomics has revealed the vast complexity of the transcriptome and we
37
38 are just beginning to understand the principles of how this translates to function,
39
40 pathophysiology and therapeutic opportunities.
41
42
43
44
45
46

47 Authors' Contributions:

48 GK, ALP, VJB, JPQ and SK conceived the idea, performed literature search, drafted
49
50 manuscript and worked with the final version. All authors participated equally.
51
52
53

54 Declaration of Conflicting Interests

55 Authors declare that they do not have conflicting interests regarding to the subject of this
56
57 manuscript.
58
59
60

Funding

ALP and SK are funded by MSWA, The Michael J. Fox Foundation, Shake It Up Australia and The Perron Institute.

For Peer Review

References

1. Mattick JS. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *BioEssays : news and reviews in molecular, cellular and developmental biology* 2003;**25**:930-9
2. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, Gingeras TR. Large-scale transcriptional activity in chromosomes 21 and 22. *Science (New York, NY)* 2002;**296**:916-9
3. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Jr., Hieter P, Vogelstein B, Kinzler KW. Characterization of the yeast transcriptome. *Cell* 1997;**88**:243-51
4. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng J-F, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen H-C, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JGR, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AFA, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang S-P, Yeh R-F, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Patrinos A, Morgan MJ, International Human Genome Sequencing C, Whitehead Institute for Biomedical Research CfGR, The Sanger C, Washington University Genome Sequencing C, Institute UDJG, Baylor College of Medicine Human Genome Sequencing C, Center RGS, Genoscope, Cnrs UMR, Department of Genome Analysis IoMB, Center GTCS, Beijing Genomics Institute/Human Genome C, Multimegabase Sequencing Center TifSB, Stanford Genome Technology C, University of Oklahoma's Advanced Center for Genome T,

1
2
3 Max Planck Institute for Molecular G, Cold Spring Harbor Laboratory LAHGC, Biotechnology
4 GBGRcf, *Genome Analysis G, Scientific management: National Human Genome Research
5 Institute USNIoH, Stanford Human Genome C, University of Washington Genome C,
6 Department of Molecular Biology KUSoM, University of Texas Southwestern Medical Center
7 at D, Office of Science USDoE, The Wellcome T. Initial sequencing and analysis of the human
8 genome. *Nature* 2001;**409**:860-921

9
10 5. Tutar Y. Pseudogenes. *Comp Funct Genomics* 2012;**2012**:424526-26

11 6. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK,
12 Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabudhe
13 NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD, Patil AH, Nanjappa V,
14 Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK,
15 Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram
16 S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K,
17 Chatterjee A, Huang TC, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar
18 S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh
19 R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TS, Hruban RH, Kerr CL, Bader GD,
20 Iacobuzio-Donahue CA, Gowda H, Pandey A. A draft map of the human proteome. *Nature*
21 2014;**509**:575-81

22 7. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T,
23 Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR,
24 Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R,
25 Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP,
26 Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B,
27 Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF,
28 Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T,
29 Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L,
30 Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J,
31 Kitamura H, Kitano H, Kollias G, Krishnan SP, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau
32 LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni
33 L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder
34 N, Nakano N, Nakauchi H, Ng P, Nilsson R, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki
35 Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF,
36 Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schonbach C,
37 Sekiguchi K, Semple CA, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D,
38 Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammaja K, Tan SL,
39 Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL,
40 Yagi K, Yamanishi H, Zabarovskiy E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale
41 RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D,
42 Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K,
43 Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano
44 K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K,
45 Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y. The transcriptional landscape
46 of the mammalian genome. *Science (New York, NY)* 2005;**309**:1559-63

47 8. Pennisi E. Genomics. ENCODE project writes eulogy for junk DNA. *Science (New York,*
48 *NY)* 2012;**337**:1159, 61

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

10. Dufva M. Introduction to microarray technology. *Methods in molecular biology (Clifton, NJ)* 2009;**529**:1-22
11. Iida K, Nishimura I. Gene expression profiling by DNA microarray technology. *Crit Rev Oral Biol Med* 2002;**13**:35-50
12. Della Beffa C, Cordero F, Calogero RA. Dissecting an alternative splicing analysis workflow for GeneChip Exon 1.0 ST Affymetrix arrays. *BMC genomics* 2008;**9**:571
13. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nature reviews Genetics* 2019;**20**:631-56
14. Hunter MC, Pozhitkov AE, Noble PA. Accurate predictions of postmortem interval using linear regression analyses of gene meter expression data. *Forensic Sci Int* 2017;**275**:90-101
15. Pardue S, Zimmerman AL, Morrison-Bogorad M. Selective postmortem degradation of inducible heat shock protein 70 (hsp70) mRNAs in rat brain. *Cellular and molecular neurobiology* 1994;**14**:341-57
16. Beach TG. Alzheimer's disease and the "Valley Of Death": not enough guidance from human brain tissue? *J Alzheimers Dis* 2013;**33 Suppl 1**:S219-33
17. Wu C, Bendriem RM, Garamszegi SP, Song L, Lee CT. RNA sequencing in post-mortem human brains of neuropsychiatric disorders. *Psychiatry Clin Neurosci* 2017;**71**:663-72
18. Koks S, Soomets U, Paya-Cano JL, Fernandes C, Luuk H, Plaas M, Terasmaa A, Tillmann V, Noormets K, Vasar E, Schalkwyk LC. Wfs1 gene deletion causes growth retardation in mice and interferes with the growth hormone pathway. *Physiological genomics* 2009;**37**:249-59
19. Koks S, Soomets U, Plaas M, Terasmaa A, Noormets K, Tillmann V, Vasar E, Fernandes C, Schalkwyk LC. Hypothalamic gene expression profile indicates a reduction in G protein signaling in the Wfs1 mutant mice. *Physiological genomics* 2011;**43**:1351-8
20. Ivask M, Pajusalu S, Reimann E, Koks S. Hippocampus and Hypothalamus RNA-sequencing of WFS1-deficient Mice. *Neuroscience* 2018;**374**:91-103
21. Schalkwyk LC, Fernandes C, Nash MW, Kurrikoff K, Vasar E, Koks S. Interpretation of knockout experiments: the congenic footprint. *Genes Brain Behav* 2007;**6**:299-303
22. Koks S, Fernandes C, Kurrikoff K, Vasar E, Schalkwyk LC. Gene expression profiling reveals upregulation of Tlr4 receptors in Cckb receptor deficient mice. *Behav Brain Res* 2008;**188**:62-70
23. Ho XD, Phung P, V QL, V HN, Reimann E, Prans E, Koks G, Maasalu K, Le NT, L HT, H GN, Martson A, Koks S. Whole transcriptome analysis identifies differentially regulated networks between osteosarcoma and normal bone samples. *Exp Biol Med (Maywood)* 2017;**242**:1802-11
24. Planken A, Kurvits L, Reimann E, Kadastik-Eerme L, Kingo K, Koks S, Taba P. Looking beyond the brain to improve the pathogenic understanding of Parkinson's disease: implications of whole transcriptome profiling of Patients' skin. *BMC Neurol* 2017;**17**:6
25. Kurvits L, Reimann E, Kadastik-Eerme L, Truu L, Kingo K, Erm T, Koks S, Taba P, Planken A. Serum Amyloid Alpha Is Downregulated in Peripheral Tissues of Parkinson's Disease Patients. *Front Neurosci* 2019;**13**:13
26. Billingsley KJ, Lattekivi F, Planken A, Reimann E, Kurvits L, Kadastik-Eerme L, Kasterpalu KM, Bubb VJ, Quinn JP, Koks S, Taba P. Analysis of repetitive element expression in the blood and skin of patients with Parkinson's disease identifies differential expression of satellite elements. *Scientific reports* 2019;**9**:4369

- 1
 - 2
 - 3
 - 4
 - 5
 - 6
 - 7
 - 8
 - 9
 - 10
 - 11
 - 12
 - 13
 - 14
 - 15
 - 16
 - 17
 - 18
 - 19
 - 20
 - 21
 - 22
 - 23
 - 24
 - 25
 - 26
 - 27
 - 28
 - 29
 - 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
27. Galichon P, Xu-Dubois YC, Buob D, Tinel C, Anglicheau D, Benbouzid S, Dahan K, Ouali N, Hertig A, Brocheriou I, Rondeau E. Urinary transcriptomics reveals patterns associated with subclinical injury of the renal allograft. *Biomarkers in medicine* 2018;**12**:427-38
28. Sole C, Goicoechea I, Goni A, Schramm M, Armesto M, Arestin M, Manterola L, Tellaetxe M, Alberdi A, Nogueira L, Roumiguie M, Lopez JI, Sanz Jaka JP, Urruticoechea A, Vergara I, Loizaga-Iriarte A, Unda M, Carracedo A, Malavaud B, Lawrie CH. The Urinary Transcriptome as a Source of Biomarkers for Prostate Cancer. *Cancers* 2020;**12**:513
29. Rhoades R, Bell DR. Medical physiology. principles for clinical medicine. Online access with subscription: LWW Health Library (Integrated Basic Science Collection). 4th ed. ed: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2013.
30. He D, Yang CX, Sahin B, Singh A, Shannon CP, Oliveria JP, Gauvreau GM, Tebbutt SJ. Whole blood vs PBMC: compartmental differences in gene expression profiling exemplified in asthma. *Allergy Asthma Clin Immunol* 2019;**15**:67
31. Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, Brown PO. Individuality and variation in gene expression patterns in human blood. *Proceedings of the National Academy of Sciences of the United States of America* 2003;**100**:1896-901
32. Sourvinou IS, Markou A, Lianidou ES. Quantification of circulating miRNAs in plasma: effect of preanalytical and analytical parameters on their isolation and stability. *J Mol Diagn* 2013;**15**:827-34
33. Jones W, Greytak S, Odeh H, Guan P, Powers J, Bavarva J, Moore HM. Deleterious effects of formalin-fixation and delays to fixation on RNA and miRNA-Seq profiles. *Scientific reports* 2019;**9**:6980
34. Buschmann D, Haberberger A, Kirchner B, Spornraft M, Riedmaier I, Schelling G, Pfaffl MW. Toward reliable biomarker signatures in the age of liquid biopsies - how to standardize the small RNA-Seq workflow. *Nucleic acids research* 2016;**44**:5995-6018
35. Carulli JP, Artinger M, Swain PM, Root CD, Chee L, Tulig C, Guerin J, Osborne M, Stein G, Lian J, Lomedico PT. High throughput analysis of differential gene expression. *J Cell Biochem Suppl* 1998;**30-31**:286-96
36. Gress TM, Wallrapp C, Frohme M, Muller-Pillasch F, Lacher U, Friess H, Buchler M, Adler G, Hoheisel JD. Identification of genes with specific expression in pancreatic cancer by cDNA representational difference analysis. *Genes Chromosomes Cancer* 1997;**19**:97-103
37. Hubank M, Schatz DG. Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic acids research* 1994;**22**:5640-8
38. Koks S, Luuk H, Nelovkov A, Areda T, Vasar E. A screen for genes induced in the amygdaloid area during cat odor exposure. *Genes Brain Behav* 2004;**3**:80-9
39. Bowler LD. Representational difference analysis of cDNA. *Methods Mol Med* 2004;**94**:49-66
40. Nelovkov A, Philips MA, Koks S, Vasar E. Rats with low exploratory activity in the elevated plus-maze have the increased expression of limbic system-associated membrane protein gene in the periaqueductal grey. *Neuroscience letters* 2003;**352**:179-82
41. Smith TP, Grosse WM, Freking BA, Roberts AJ, Stone RT, Casas E, Wray JE, White J, Cho J, Fahrenkrug SC, Bennett GL, Heaton MP, Laegreid WW, Rohrer GA, Chitko-McKown CG, Perteu G, Holt I, Karamycheva S, Liang F, Quackenbush J, Keele JW. Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle. *Genome research* 2001;**11**:626-30
42. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-

- 1
2
3 Dale AL, Brown PO, Botstein D. Molecular portraits of human breast tumours. *Nature*
4 2000;**406**:747-52
- 5
6 43. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van
7 de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE,
8 Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor
9 subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the*
10 *United States of America* 2001;**98**:10869-74
- 11
12 44. Ho XD, Nguyen HG, Trinh LH, Reimann E, Prans E, Koks G, Maasalu K, Le VQ, Nguyen
13 VH, Le NTN, Phung P, Martson A, Lattekivi F, Koks S. Analysis of the Expression of Repetitive
14 DNA Elements in Osteosarcoma. *Frontiers in genetics* 2017;**8**:193
- 15
16 45. Schwaederle M, Zhao M, Lee JJ, Lazar V, Leyland-Jones B, Schilsky RL, Mendelsohn J,
17 Kurzrock R. Association of Biomarker-Based Treatment Strategies With Response Rates and
18 Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. *JAMA Oncol*
19 2016;**2**:1452-59
- 20
21 46. Rodon J, Soria JC, Berger R, Miller WH, Rubin E, Kugel A, Tsimberidou A, Saintigny P,
22 Ackerstein A, Brana I, Loriot Y, Afshar M, Miller V, Wunder F, Bresson C, Martini JF, Raynaud
23 J, Mendelsohn J, Batist G, Onn A, Tabernero J, Schilsky RL, Lazar V, Lee JJ, Kurzrock R. Genomic
24 and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat Med*
25 2019;**25**:751-58
- 26
27 47. Taylor JC, Martin HC, Lise S, Broxholme J, Cazier J-B, Rimmer A, Kanapin A, Lunter G,
28 Fiddy S, Allan C, Aricescu AR, Attar M, Babbs C, Becq J, Beeson D, Bento C, Bignell P, Blair E,
29 Buckle VJ, Bull K, Cais O, Cario H, Chapel H, Copley RR, Cornall R, Craft J, Dahan K, Davenport
30 EE, Dendrou C, Devuyst O, Fenwick AL, Flint J, Fugger L, Gilbert RD, Goriely A, Green A, Greger
31 IH, Grocock R, Gruszczyk AV, Hastings R, Hatton E, Higgs D, Hill A, Holmes C, Howard M,
32 Hughes L, Humburg P, Johnson D, Karpe F, Kingsbury Z, Kini U, Knight JC, Krohn J, Lambie S,
33 Langman C, Lonie L, Luck J, McCarthy D, McGowan SJ, McMullin MF, Miller KA, Murray L,
34 Németh AH, Nesbit MA, Nutt D, Ormondroyd E, Oturai AB, Pagnamenta A, Patel SY, Percy M,
35 Petousi N, Piazza P, Piret SE, Polanco-Echeverry G, Popitsch N, Powrie F, Pugh C, Quek L,
36 Robbins PA, Robson K, Russo A, Sahgal N, van Schouwenburg PA, Schuh A, Silverman E,
37 Simmons A, Sørensen PS, Sweeney E, Taylor J, Thakker RV, Tomlinson I, Trebes A, Twigg SR,
38 Uhlig HH, Vyas P, Vyse T, Wall SA, Watkins H, Whyte MP, Witty L, Wright B, Yau C, Buck D,
39 Humphray S, Ratcliffe PJ, Bell JI, Wilkie AO, Bentley D, Donnelly P, McVean G. Factors
40 influencing success of clinical genome sequencing across a broad spectrum of disorders.
41 *Nature genetics* 2015;**47**:717-26
- 42
43 48. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, Bolduc V,
44 Waddell LB, Sandaradura SA, O'Grady GL, Estrella E, Reddy HM, Zhao F, Weisburd B,
45 Karczewski KJ, O'Donnell-Luria AH, Birnbaum D, Sarkozy A, Hu Y, Gonorazky H, Claeys K, Joshi
46 H, Bournazos A, Oates EC, Ghaoui R, Davis MR, Laing NG, Topf A, Kang PB, Beggs AH, North
47 KN, Straub V, Dowling JJ, Muntoni F, Clarke NF, Cooper ST, Bonnemann CG, MacArthur DG.
48 Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl*
49 *Med* 2017;**9**
- 50
51 49. Irigoyen N, Firth AE, Jones JD, Chung BY, Siddell SG, Brierley I. High-Resolution Analysis
52 of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathog*
53 2016;**12**:e1005473
- 54
55 50. Benson MD, Dasgupta NR, Monia BP. Inotersen (transthyretin-specific antisense
56 oligonucleotide) for treatment of transthyretin amyloidosis. *Neurodegener Dis Manag*
57 2019;**9**:25-30
- 58
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
51. Wein N, Vulin A, Findlay AR, Gumienny F, Huang N, Wilton SD, Flanigan KM. Efficient Skipping of Single Exon Duplications in DMD Patient-Derived Cell Lines Using an Antisense Oligonucleotide Approach. *J Neuromuscul Dis* 2017;**4**:199-207
52. Li D, Mastaglia FL, Fletcher S, Wilton SD. Precision Medicine through Antisense Oligonucleotide-Mediated Exon Skipping. *Trends Pharmacol Sci* 2018;**39**:982-94
53. Aung-Htut MT, McIntosh CS, Ham KA, Pitout IL, Flynn LL, Greer K, Fletcher S, Wilton SD. Systematic Approach to Developing Splice Modulating Antisense Oligonucleotides. *Int J Mol Sci* 2019;**20**

For Peer Review