Human transcriptome — molecular neurobiology

NIKO NJIRIĆ[A], FILIP ĐERKE[B]
UNIVERSITY OF ZAGREB SCHOOL OF MEDICINE

DOI: http://dx.doi.org/10.17486/gyr.3.1031

[A]0000-0002-5249-9996
[B]0000-0001-9408-9488

SUMMARY: The transcriptome consists of all RNAs transcribed from a genome. Methods of transcriptome analysis include DNA microarray, as well as RNA-sequencing. The main principle of transcriptome analysis in the study of disease is to map RNA products, from genes responsible for mutations (identified previously by genomic analysis), into a healthy human transcriptome map, and thereafter examine the spatial and temporal differences in gene product occurence, as well as their quality and quantity. This dynamic nature of transcriptomics provides a thorough insight into disease pathogenesis. By identifying co-expressed genes, commonly found in many specimens with the same comorbidity, interconnected gene products are displayed clearly, distinguishing the disease-driving mutations from the silent ones. Diseases in which vast transcriptome studies are being conducted are: Alzheimer's disease, autism spectrum disorder, schizophrenia, Tourette syndrome etc.

KEYWORDS: transcriptome, genome, neurobiology, differential expression, transcriptomics methods

The transcriptome consists of all RNAs transcribed from a genome.[1] There are qualitative and quantitative differences in RNA molecules a cell produces during its lifespan. Cells belonging to different tissues, and even very small different subunits of the same tissue, can vary drastically in RNAs they produce. This means that cells, tissues and organisms can be analyzed from a standpoint concerning the RNA molecules they produce at a given time and space. Use of this approach, opposed to the relatively fixed nature of genomic analyses, gives a dynamic, spatio-temporal insight into normal human development, as well as development of disease.[2]

Methods used in transcriptome analysis are: DNA microarray, which requires a pre-existing genome and transcriptome map as a reference, next generation sequencing (RNA-seq), which enables de novo transcriptome assembly. Real-time quantitative PCR (qRT-PCR) is used for secondary validation of previously extracted RNA.[3]

Neurobiology uses transcriptomics to study differential gene expression and alternative splicing of genes in the central nervous system during development and adulthood, with their aberrations, responsible for psychiatric and neurologal disorders.[4]

Methods

In order to study changes in transcriptome of individuals with central nervous system abnormalities, interindividual differences and similarities of healthy controls need to be established first, taking age and sex into account. Clinically remarkable specimens can thereafter be compared to assess the differences. There are 2 transcriptome profiling methods: microarray and RNA-Seq.

Microarray is a method in which fluorescently labeled cDNA is hybridized with microchips containing many specific DNA sequences, and scanned for results of hybridization.[5]

A study conducted by Kang et al. in 2011 serves as an example of transcriptome analysis using DNA microarray. 57 specimens, ranging from 5.7 postconception weeks (PCW) to 82 years of age were divided into 15 age groups. All specimens were histologically confirmed to be disease free. In first two age groups, 4-8 PCW and 8-10 PCW, 8 and 10 brain regions were dissected for analysis, respectively. In the other 13 age groups there were 16 regions of interest prepared for analysis: cerebellum, mediodorsal thalamic nucleus, striatum, amygdala, hippocampus, and 11 neocortical areas. After dissection, RNA was extracted from the brain tissue and reversely transcribed to generate cDNA. It was then hybridized with an exon array containing 1.4 million DNA sequences (probe sets), and scanned for results.[3]

RNA-seq is a newer method compared to DNA microarray. Its main advantage is that it does not rely on previously designed probes which allows for novel transcripts and isoforms to be discovered. It also has a wider dynamic range, meaning it has a higher sensitivity in detecting genes with very low or very high expression.[6] The downside, however, is its higher price.

Firstly, the RNA molecules are reversely transcribed to cDNA, which can then optionally be fragmented. The next step is the attachment of adaptor sequences to one or both ends of cDNA fragments. These sequences are specific for each RNA-sequencing platform such as Illumina IG18–21, 23,24, Applied Biosystems SOLiD22 and Roche 454 Life Science26–28 systems. The cDNA fragments binded to the adapters are then optionally amplified and sequenced on a sequencing platform.[7]

Each platform has its own mechanism of action, based on the usage of DNA polymerase (sequencing-by-synthesis) or DNA lygase (sequencing-by-ligation).[8] Results can be alligned to a reference genome or transcription profile, or a genome-scale transcription map can be assembled de novo.[7]

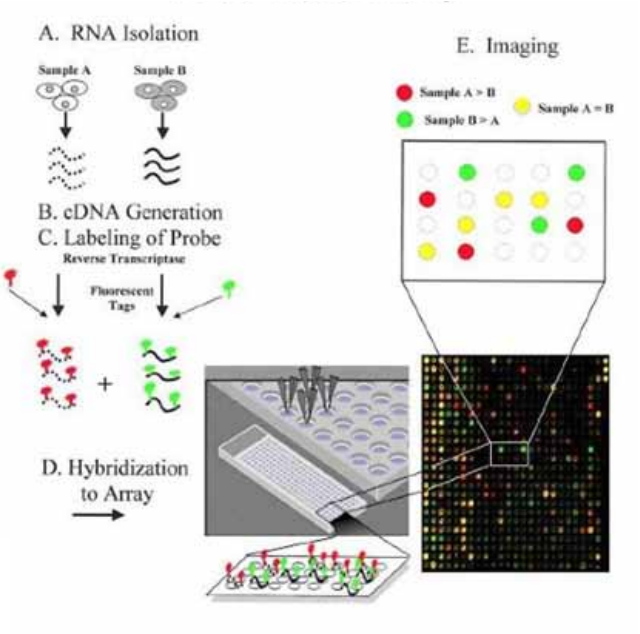

Fig 1. DNA microarray: RNA is isolated (A), reversely transcribed to generate cDNA (B), labeled fluorescently (C), hybridized to array (D) and scanned for results - red means the gene is active in sample A only, green means the gene is active in sample B only, and yellow means it is active in both samples (E) (image from:https://nanohub.org/resources/17701/watch?resid=17812)

Application of transcriptome analysis

The main principle of transcriptome analysis in the study of disease is to map RNA products, from genes responsible for mutations (identified previously by genomic analysis), onto a healthy human transcriptome map, and thereafter examine the spatial and temporal differences in gene product occurence, as well as their quality and quantity.
Diseases in which vast transcriptome studies are being conducted are: Alzheimer's disease, autism spectrum disorder, schizophrenia, Tourette syndrome etc.

Alzheimer's disease

Alzheimer's disease is the most common form of dementia. It is characterized by debilitating and progressive episodic memory loss, diffuculty with language and decision making, and loss of motor control, incontinence, and mutism. Regarding age of disease onset, Alzheimer's disease (AD) is divided into two types: familial early-onset Alzheimer's disease (FEOAD), which represents 5% of total AD patients, while late-onset Alzheimer's disease (LOAD) covers the remaining 95%. In FEOAD, genomic analyses concluded mutations in *APP, PSEN1* and *PSEN2* are responsible for accumulation of amyloid-beta plaques. Mutations in the *APP* and *PSEN1* genes explain the majority of FEOAD cases, while mutations in the *PSEN2* are singnificantly rarer with most recorded mutations in this gene arising within a single pedigree known as the Volga Germans. In an attempt to decypher the more complex, LOAD pathogenesis, brain transcriptome studies have been conducted. Two genes are reported to have exhibited the strongest correlation with disease occurence when overexpressed, *TYROBP* and *APOE*.[9] *TYROBP* is expressed in microglial cells and has a role in amyloid-beta turnover and neuronal damage.[10] *APOE* is expressed in astrocytes and the APOE4 phenotype holders exhibit GABA interneuron dysfunction which is thought to have implications on disease pathogenesis.[10] Additionally, a recent sequencing experiment involving over 1700 Icelandic participants has revealed a rare coding mutation in *APP* that is not only protective for AD, but also appears to be associated with generally reduced symptoms of cognitive decline in aging carriers. While the individual effect size of any one of these non-APOE varriants is much smaller than that of *APOE*, researchers still hope the combined genotypic knowledge of these susceptibility genes may lead to better genetic testing for AD.

Autism spectrum disorder

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by behavioral deficits, restricted interests and stereotypic behavior, affecting 1 in 68 children.[9] The age of onset is from 3 years of age to early childhood.[9] Although, with increasing recognition of autism, it may also be diagnosed in adults who have received other diagnoses in the past (e.g. Intellectual Disability). Dysfunction in synapse development is thought to be the main factor responsible for ASD. There are as many as 1000 genes suggested to have involvement in ASD pathogenesis which points to a heterogenous nature of ASD development. This calls for a genome-wide transcriptome analysis to assess the temporal and spatial character of converging pathways, leading to disease formation.[11] Many loci have been imlicated through various genome-wide association studies, but in some cases, the findings hav been contradictory among different studies. A study by Wilsey et al. focused on 9 high confidence ASD genes: *ANK2, CHD8, CUL3, DYRK1A, GRIN2B, KATNAL2, POGZ, SCN2A* and *TBR1* and constructed their co-expression networks. The results implied the site of convergence is localised in glutamatergic projection neurons in layers 5 and 6 of prefrontal and motor-somatosensory cortex in periods from 10-24 postconception weeks, pointing to a specific region of interest, regarding time and space, for future pathophysiological studies.[12]

Schizophrenia

Schizophrenia (SCZ) is a psychiatric disorder characterized by deviations in thought processes and perception of reality. Its genetic background still remains to be fully elucidated. There are no agreed upon monogenic forms, no single gene has been demonstrated to cause SCZ. Strongest risk for SCZ is a positive family history. Advanced paternal age is associated with elevated risk. De novo single-nucleotide polymorphism (SNP) and copy-number variants (CNV) rates are increased in SCZ. Significant genome-wide associations exist for both CNVs and SNPs. Large CNVs are associated with larger risk, but produce pleiotropic phenotypes (ie, not specific for SCZ). Testing for CNVs or SNPs is not yet clinically validated for diagnosis or treatment selection. Gulsuner et al. have conducted a study on affected individuals from otherwise healthy families. After genomic analyses of mutations, they mapped the responsible genes onto healthy human transcriptome profiles. The most

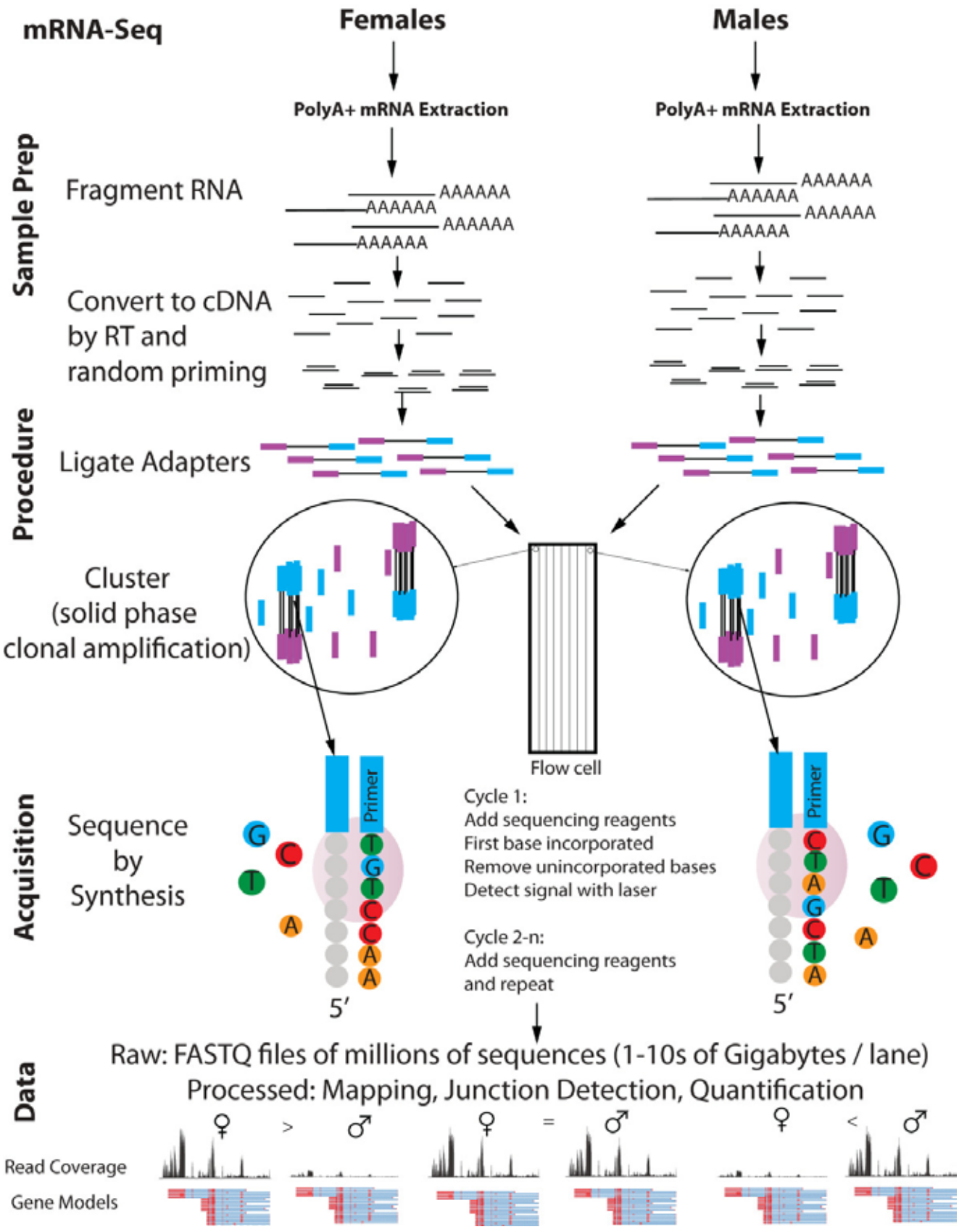

Fig 2. RNA-Seq: The RNAs are fragmented, reversely transcribed to cDNA, attached to adaptor sequences, amplified and then sequenced on a platform which uses sequencing-by-synthesis. (image from:http://www.biomedcentral.com/1741-7007/9/34/figure/F2?highres=y)

remarkable finding was significant convergence of coexpressed genes and protein-protein interactions in the dorsolateral and ventrolateral prefrontal cortex during fetal development (10-26 post-conceptual weeks), establishing a point in time and space critical for schizophrenia susceptibility.[13]

Tourette syndrome

Tourettes syndrome (TS) is a disorder characterized by young age of onset, as well as motor and vocal tics. Transcriptome analysis of TS patients' striatum, performed by Lennington et al. provided insight into 2 diferentially expressed gene groups. The downregulated one was located in the projection neurons of the striatum and comprised of genes responsible for acethylcholine synthesis and transport, cholinergic receptors and homeodomain genes. This decrease of neuronal signaling is thought to be involved in the pathogenesis of TS. The upregulated group of genes was comprised of immune related genes, active in microglial cells of the striatum, supporting the immunological hypothesis of TS development.[14]

Conclusion

Transcriptomics are becoming widely used in many areas of neurobiology. Neurological, psychiatric and neurooncological disorders are becoming increasingly understood thanks to spatial and temporal mapping of the human brain suffering from some form of disease. Many diseases with a heterogenous mutation profile will benefit from transcriptome profiling, as it will help to identify a convergence point of all aberrantly expressed genes, and elucidate disease pathways.

LITERATURE:
1. ERIC R. KANDEL, JAMES H. SCHWARTZ, THOMAS M. JESSELL, STEVEN A. SIEGELBAUM, A. J. HUDSPETH. PRINCIPLES OF NEURAL SCIENCE, FIFTH EDITION; 2013.
2. PEĆINA-ŠLAUS N, PEĆINA M. ONLY ONE HEALTH, AND SO MANY OMICS. CANCER CELL INT 2015; 15:64.
3. KANG HJ ET AL. SPATIO-TEMPORAL TRANSCRIPTOME OF THE HUMAN BRAIN. NATURE. 2011; 478:483-489.
4. JOHNSON MB ET AL. FUNCTIONAL AND EVOLUTIONARY INSIGHTS INTO HUMAN BRAIN DEVELOPMENT THROUGH GLOBAL TRANSCRIPTOME ANALYSIS. NEURON. 2009; 62:494–509.
5. LASHKARI DA ET AL. YEAST MICROARRAYS FOR GENOME WIDE PARALLEL GENETIC AND GENE EXPRESSION ANALYSIS. PROC NATL ACAD SCI USA. 1997; 94:13057-13062.
6. ZHAO S, FUNG-LEUNG WP, BITTNER A, NGO K, LIU X. COMPARISON OF RNA-SEQ AND MICROARRAY IN TRANSCRIPTOME PROFILING OF ACTIVATED T CELLS. PLOS ONE 9(1): E78644. DOI:10.1371/JOURNAL.PONE.0078644
7. WANG G, GERSTEIN M, SNYDER M. RNA-SEQ: A REVOLUTIONARY TOOL FOR TRANSCRIPTOMICS. NAT REV GENET. 2009; 10(1): 57–63.
8. SUN HJ, CHEN J, NI B, YANG X, WU YZ. RECENT ADVANCES AND CURRENT ISSUES IN SINGLE-CELL SEQUENCING OF TUMORS. CANCER LETT. 2015; 365: 1–10.
9. HAN G, SUN J, WANG J, BAI Z, SONG F, LEI H. GENOMICS IN NEUROLOGICAL DISORDERS. GENOMICS PROTEOMICS BIOINFORMATICS. 2014; 12:156–163.
10. ZHANG B ET AL. INTEGRATED SYSTEMS APPROACH IDENTIFIES GENETIC NODES AND NETWORKS IN LATE-ONSET ALZHEIMER'S DISEASE. CELL. 2013; 153:707–720.
11. PARIKSHAK NN, LUO R, ZHANG A, ET AL. INTEGRATIVE FUNCTIONAL GENOMIC ANALYSES IMPLICATE SPECIFIC MOLECULAR PATHWAYS AND CIRCUITS IN AUTISM. CELL. 2013; 155:1008–1021.
12. WILLSEY AJ. COEXPRESSION NETWORKS IMPLICATE HUMAN MIDFETAL DEEP CORTICAL PROJECTION NEURONS IN THE PATHOGENESIS OF AUTISM. CELL. 2013; 155: 997–1007.
13. GULSUNER S ET AL. SPATIAL AND TEMPORAL MAPPING OF DE NOVO MUTATIONS IN SCHIZOPHRENIA TO A FETAL PREFRONTAL CORTICAL NETWORK. CELL. 2013; 154:518–529.
14. LENNINGTON JB ET AL. TRANSCRIPTOME ANALYSIS OF THE HUMAN STRIATUM IN TOURETTE SYNDROME. BIOL. PSYCHIATRY. AVAILABLE ONLINE 24 JULY 2014, DOI:10.1016/J.BIOPSYCH.2014.07.018

Ljudski transkriptom- molekularna neurobiologija

SAŽETAK: Transkriptom sadrži sve RNA molekule dobivene transkripcijom iz genoma. Metode analize uključuju DNA-mikropostroj tehnologiju i RNA sekvencioniranje. Glavno načelo analize transkriptoma u proučavanju bolesti je mapiranje RNA produkata gena odgovornih za mutacije (prethodno identificiranih pomoću analize genoma) na transkriptom zdravog čovjeka i posljedično proučavanje vremensko-prostornih razlika u pojavljivanju produkata gena, kao i njihovoj kvaliteti i kvantiteti. Dinamična priroda transkriptomike omogućava detaljan uvid u patogenezu bolesti. Identificiravši ko-eksprimirane gene pronađene u velikom broju pacijenata sa istim komorbiditetom, međusobno povezani produkti gena se jasnu prikazuju, samim time omogućavajući razlučivanje mutacija odgovornih za nastanak bolesti, od onih nebitnih. Primjeri bolesti na kojima se trenutno vode studije transkriptoma su autizam, shizofrenija, Touretteov sindrom i Alzheimerova bolest.

KLJUČNE RIJEČI: transkriptom, genom, neurobiologija, diferencijalna ekspresija, metode u transkriptomici