Analysis of Forensic Transcriptomes by Massively Parallel Sequencing

Cordula Haas*

Department of Forensic Genetics, University of Zurich, Zurich, Switzerland

Introduction

Over the past ten years, the field of forensic genetics' use of transcriptome analyses has grown tremendously. The earliest examinations and fundamental applications were body liquid and tissue recognizable proof, utilizing designated RNA records and a converse record endpoint PCR strategy. The method has been successfully utilized in casework and a number of markers have been identified for the body fluids and tissues that are most relevant to forensics. The development of Massively Parallel Sequencing (MPS) gave rise to brand-new viewpoints and presented new chances to advance the field. In contrast to autosomal DNA, which contains two copies in a cell, abundant RNA species are expressed in high copy numbers. It was demonstrated that whole transcriptome sequencing (RNA-Seq) of postmortem and forensically relevant body fluids is possible. An overview of forensic transcriptome analyses and their applications is provided in this review. Both targeted MPS methods and the entire transcriptome are included in the methods. Body fluid/tissue identification, determining the age of stains and the donor, estimating the post-mortem interval, and post-mortem death investigations all benefit from high-resolution forensic transcriptome analyses with MPS [1].

Transcriptomics and its potential application in law enforcement Bauer published a review titled "RNA in forensic science" in 2008, just as RNA analysis had only recently emerged in forensic genetics. He anticipated the significant potential of RNA analyses to resolve forensic issues, such as the identification of body fluid, the estimation of the age of wounds, the estimation of the post-mortem interval (PMI), the estimation of the age of stains, and the identification of the cause of death. Numerous RNA-based studies have since been published on these subjects, particularly those pertaining to body fluid identification applications. By distinguishing menstrual blood from peripheral blood on suspected perpetrators in sexual assault cases, it can be important to support sexual versus social intercourse by determining the cellular (pheno) type from which the DNA originated [2].

Description

In the last two decades, Massively Parallel Sequencing (MPS) has revolutionized research and diagnostics by reducing operational costs and increasing throughput. It is now possible to study transcriptomes or entire genomes. The complete set of RNA transcripts—both coding and non-coding—in a particular cell, cell population, or tissue is referred to as the transcriptome in this review. In addition, there are methods for concentrating on specific RNA biomarkers and subgenomic DNA regions. The vast amount of data being generated necessitates sophisticated bioinformatics expertise and appropriate analysis pipelines. Even though forensic samples pose a real obstacle to any MPS experiment due to their degraded nature and low number of nucleic acid species, a number of studies have demonstrated promising results when analyzing these kinds of samples. Additionally, RNA-Seq, or whole transcriptome sequencing, was successfully utilized on forensic samples [3].

For the contextualization of various stains, determining the origin of a biological trace is crucial, and as a result, it may provide crucial information to authorities enforcing the law. Presumptive tests like chemical, immunological, and protein catalytic activity tests, as well as spectroscopic methods and microscopy, are frequently used to identify body fluids. These tests are not available for all forensically relevant body fluids (such as vaginal secretion), and it is difficult to distinguish between related body fluids like blood and menstrual blood due to cross-reactivity with other substances or tissues. With RNA profiling, most of the common forensically relevant body fluids and tissues found in dried stains can now be identified with high certainty, including blood, semen, saliva, vaginal secretions, menstrual blood, and skin. In the past, PCR/CE methods were utilized, but MPS methods have increasingly emerged in recent years [1].

Zubakov et al. conducted a proof-of-principle study. Presented a method for the simultaneous use of the Ion Torrent PGM and parallel targeted DNA/RNA sequencing to analyze forensic mRNAs, amelogenin, and STRs. The measure included 9 autosomal STRs, the AMELX/AMELY framework for sex recognizable proof and 12 mRNA markers, involving separate work processes for RNA and DNA examination. STR sequencing analysis of the tissue sample donors was 100% consistent with conventional STR profiling, and all samples from all forensically relevant tissues tested were able to be identified by mRNA alone. As a result, it was demonstrated that it was possible to simultaneously sequence a variety of nucleic acid markers, such as STRs, amelogenin insertion/deletion, and mRNAs, for a variety of forensic purposes, including the identification of individuals, sexes, and tissues. Lin and co. used MPS data from both fresh and degraded body fluids to find "transcript stable regions" or "StaRs" in target transcripts with high read coverage and particularly stable sequences [4].

A novel idea they propose is that StaR-targeted primers can consistently and precisely amplify a wide range of RNA biomarkers in a variety of body fluids with varying degradation levels. Hanson and co. created a targeted RNA sequencing assay for body fluid identification. The 33 markers in the assay, which was made for Illumina MiSeq/FGx platforms, can be used to identify blood, menstrual blood, sperm, saliva, vaginal secretions, and skin. The assay's sensitivity and specificity were checked, and it worked on both single-source and mixed stains. In order to make inferences about the body fluid or tissue, we looked into two classification methods: the percentage of reads in a sample that are due to each of the six body fluids or tissues that were tested and the differential gene expression between samples that was found by agglomerative hierarchical clustering. Dorum and co. built a probabilistic model based on partial least squares (PLS) and linear discriminant analysis (LDA) using this MiSeq data set to predict a stain's origin. Improved predictions were achieved as a result of the model's incorporation of quantitative data (MPS read counts) rather than merely the presence or absence of markers. Additionally, the model was successfully used to identify the various components of mixed body fluid samples [5].

Conclusion

To determine whether targeted mRNA sequencing can be used to identify bodily fluids, a collaborative exercise was planned within the EUROFORGEN/ EDNAP laboratories. An in-house assay for the Ion Torrent PGM/S5 platform with 29 markers and the aforementioned Illumina MiSeq/FGx assay were evaluated. There was some between research facility fluctuation in read counts, yet generally the aftereffects of the labs were comparable. With low input samples, the in-house PGM/S5 workflow appeared less reliable. An MPS assay for body fluid identification based on the Ion S5 system was recently introduced. According to the prototype assays that have been presented, a promising method for identifying body fluids is the analysis of mRNA by targeting tissue- or body fluid-specific amplicons. Improvements in marker selection, library preparation, and sequencing will be based on the corresponding results.

Acknowledgement

None.

Conflict of Interest

None.

References

1. Tolbert, Mariah, Sheree J. Finley, Silvia D. Visonà and Shivani Soni, et al. "The thanatotranscriptome: Gene expression of male reproductive organs after death." *Gene* 675 (2018): 191-196.

2. Hunter, M. Colby, Alex E. Pozhitkov and Peter A. Noble. "Accurate predictions of postmortem interval using linear regression analyses of gene meter expression data." *Forensic Sci Int* 275 (2017): 90-101.

3. Zhu, Yizhang, Likun Wang, Yuxin Yin and Ence Yang. "Systematic analysis of gene expression patterns associated with postmortem interval in human tissues." *Sci Rep* 7 (2017): 1-12.

4. Neubauer, Jacqueline, Maria Rita Lecca, Giancarlo Russo and Christine Bartsch, et al. "Exome analysis in 34 sudden unexplained death (SUD) victims mainly identified variants in channelopathy-associated genes." *Int J Legal Med* 132 (2018): 1057-1065.

5. Gaertner, Anna, Patrick Schwientek, Peter Ellinghaus and Holger Summer, et al. "Myocardial transcriptome analysis of human arrhythmogenic right ventricular cardiomyopathy." *Physiol Genomics* 44 (2012): 99-109.