# NGS-SSR based Biomarkers for Diagnosis of Human Diseases

Arghavan Alisoltani

*OMICS Research Group, Department of Biotechnology, Faculty of Applied and Computer Sciences, Vaal University of Technology, South Africa.*

*arghavana1@vut.ac.za*

*arghavana85@gmail.com*

Anousheh Alisoltani

*School of veterinary sciences, University of Shahrekord, Iran*

*nusha.soltani@gmail.com*

Esmaeil Ebrahimi

*The University of Adelaide, Australia*

*ebrahimiet@gmail.com*

Hossein Fallahi

*The Razi University of Kermanshah*

*fallahi.hossein@gmail.com*

Naser Aliye Feto

*OMICS Research Group, Department of Biotechnology, Faculty of Applied and Computer Sciences, Vaal University of Technology, South Africa.*

*naserf@vut.ac.za*

**Abstract**

**Background**: NGS data can be used to investigate the relationship between SSRs and human diseases to uncover the influence of SSRs on the development and progression of diseases. These insights can then be translated to clinical benefits, including development of reliable biomarkers and effective strategies for prevention and therapy. So far, there has been no report on the application of NGS-SSR in cancer diagnosis or other human diseases. One of the main goals of this study was development of new method for discovery of reliable biomarkers. In this case study, we focused on identification of biomarkers shared across variety of cancers.

**Methodology**: In brief, the different types of datasets were selected to SSR analysis. NGS data were checked for quality and trimmed before SSR analysis. The sequences for each healthy and cancerous samples were analyzed for SSR motifs using SSR locator software. Then to obtain differentially expressed (DE) SSRs, the number of SSRs was compared between cancerous and healthy libraries. To find target genes with DE SSRs, sequences with specific altered SSRs were further extracted and annotated. In the final step, Gene Ontology of target genes were performed.

**Results and Discussion**: In this study, for the first time, we developed NGS-SSR analysis pipeline to find reliable biomarkers for cancer diagnosis. Analysis of RNA-seq data revealed a significant alternation in the frequency of SSR motifs during cancer progression. Among different subtypes of SSR motifs, trinucleotide repeats are of great interest because of their role in many human diseases such as cancer. RNA-seq SSR showed an increase in the frequency of GCC/GGC and GCG/CGC motifs in cancerous tissues compared to healthy ones. Based on the observed results, we suggest the potential use of mentioned motifs as biomarkers in cancer diagnosis. In conclusion, NGS-SSRs can be reliable biomarkers for diagnosis of human diseases. With decreased cost of RNA-Seq in recent years, it can be feasible to perform RNA-seq SSR for each patient. In addition, this technique has potential to directly apply on raw data, which decreases the need for complex computational analysis.

**Key words:** Biomarker, Cancer, Diagnosis, NGS-SSR, RNA-seq