

InFusion: advancing discovery of fusions genes and chimeric transcripts from RNA-sequencing data

Konstantin Okonechnikov¹, Aki Imai-Matsushima², Lukas Paul³, Alexander Seitz³, Thomas F. Meyer², Fernando Garcia-Alcalde⁴

¹ Division of Pediatric Neurooncology, German Cancer Research Center, Heidelberg, Germany

² Department of Molecular Biology, Max Planck Institute for Infection Biology, Berlin, Germany

³ Lexogen GmbH, Campus Vienna Biocenter 5, Vienna, Austria

⁴ Roche Pharma Research and Early Development, Infectious Diseases, Roche Innovation Center Basel, Switzerland

Functional fusion genes and chimeric transcripts have been shown to occur in cancers due to genomic rearrangements as well as in non-cancerous cells due to trans-splicing or transcriptome machinery failure. Correct and detailed detection of fusions is important in scientific research and in precision cancer medicine as well. RNA sequencing appeared to be an effective method for discovery of fusions. We have designed and implemented a novel toolkit called InFusion for chimeric transcript discovery from RNA-seq data. In comparison to other existing methods, our approach introduces several unique features such as discovery of fusions involving intergenic regions and detection of anti-sense chimeras based on the strand-specificity of the sequencing library. Additionally, the toolkit includes several advanced post-analysis steps such as comparison of results among well-known existing tools and design of sequences for further experimental validation.

Using simulated and public data we demonstrated that InFusion has superior detection sensitivity compared to other existing methods and is able to discover a wider spectrum of fusion events that can occur in the transcriptome. Importantly, we also performed deep RNA sequencing of two prostate cancer cell lines. From this experimental data analysis we discovered in-silico and verified in-vitro 26 novel fusion events, including alternatively spliced fusion isoforms and chimeric RNAs involving non-exonic regions. Moreover, we confirmed four fusions that involve intergenic regions. To our knowledge, discovery of such events has not been addressed previously, despite their potential to encode functional proteins or regulate gene transcription.

The detailed landscape of the chimeric RNAs, mechanisms underlying their genesis and their functional roles are yet to be studied. InFusion may prove to be a useful tool for detecting the whole scope of possible events. The manuscript describing the method is published in PLOS One and the open-source software toolkit is available for download at: <http://bitbucket.org/kokonech/infusion>

Manuscript reference: Okonechnikov, K., et al. *PloS one* 11.12 (2016): e0167417.