BITS :: Call for Abstracts 2023 - Oral communication

Type	Oral communication
Session	Genomics, transcriptomics, epigenomics and epitranscriptomics
Title	IsoPrimer: a primer-designing pipeline for amplicon-based validation of genes and splicing variants expression.
All Authors	Filomena E(1), Picardi E(1,2) , Pesole G(1,2), D'Erchia AM(1,2)

Affiliation

(1) Department of Biosciences, Biotechnologies and Environment, University of Bari A. Moro, Bari Italy
(2) Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies, CNR, Bari

Motivation

Genes perform multiple functions by generating multiple transcripts through alternative splicing. The accurate quantification of transcript variants in different conditions is crucial for their functional characterization and requires accurate design of PCR primer pairs that may selectively target specific isoforms. Primer design may result tricky depending on the complexity of the splicing pattern both for detecting specific isoforms and for quantifying the overall gene expression. Whilst several primer designing tools currently exist, an all-in-one free and open source software solution capable of comprehensively tackling the problem is not available. Therefore, we developed the IsoPrimer pipeline which automates the designing process of primer pairs for PCR, aimed at addressing both issues above.

Methods

IsoPrimer is written in R (version 3.6.1) and Bash (4.2.46(2)). The pipeline leverages Kallisto (1), Primer3 (2) and EMBOSS PrimerSearch (3), respectively to: 1) assess the expression level of the different isoforms of genes of interest from RNA-seq data; 2) design all primer pairs overlapping exon-exon junctions in the expressed isoforms; 3) verify the specificity of the designed primer pairs. In the first step, Kallisto is run on the RNA-seq samples specified in the sample list indicated by the user. Kallisto output files are manipulated to compile a transcript-level TPM matrix which is used to assess the expression of isoforms. To retrieve gene information, such as ENSEMBL ID, official gene symbol and position of exon-exon junctions in splicing variants, the ENSEMBL annotation and transcriptome provided by the user are parsed to build an information dense table used to quickly retrieve, via vectorized R functions, the nucleotidic position of each splicing junction and instruct Primer3 to design primer pairs, with at least one splicing-specific primer, compliant to the constraints defined by the user via the Primer3 boulder.IO configuration file. By leveraging the capabilities of the "parallel" R package, the specificity of the unique primer pairs designed for the expressed isoforms of target genes is simultaneously checked by performing in-silico PCRs with PrimerSearch. IsoPrimer finally computes a score value for each primer pair, considering the ability to proxy the global gene expression and providing information on the represented transcript isoforms.

Results

IsoPrimer is written in R (version 3.6.1) and Bash (4.2.46(2)). IsoPrimer was designed as a modular pipeline capable of automatically performing all steps necessary to design primer pairs for specific splicing variants. The pipeline leverages Kallisto (1), Primer3 (2) and EMBOSS PrimerSearch (3), respectively to: 1) assess the expression level of the different isoforms of genes of interest from RNA-seq data; 2) design all primer pairs overlapping exon-exon junctions in the expressed isoforms; 3) verify the specificity of the designed primer pairs. In the first step, Kallisto is run on the RNA-seq samples specified in the sample list indicated by the user. Kallisto output files are manipulated to compile a transcript-level TPM matrix which is used to assess the expression of isoforms. To retrieve gene information, such as ENSEMBL ID, official gene symbol and position of exon-exon junctions in splicing variants, the ENSEMBL annotation and transcriptome provided by the user are parsed to build an information dense table used to quickly retrieve, via vectorized R functions, the nucleotidic position of each splicing junction and instruct Primer3 to design primer pairs, with at least one splicing-specific primer, compliant to the constraints defined by the user via the Primer3 boulder.IO configuration file. By leveraging the capabilities of the "parallel" R package, the specificity of the unique primer pairs designed for the expressed isoforms of target genes is simultaneously checked by performing in-silico PCRs with PrimerSearch. IsoPrimer finally computes a score value for each primer pair, considering the ability to proxy the global gene expression and providing information on the represented transcript isoforms. We believe IsoPrimer may represent a valuable contribution to transcriptomics analyses aiming at validating the expression profile of genes of interest, resulting deregulated in RNA-seq analysis. Compared to widely used web tools as PrimerBlast (4), IsoPrimer allows the user to specifically focus on expressed splicing variants of a gene of interest, to simultaneously design primers for multiple target genes and isoforms and to run specificity checks of the designed primer pairs on the transcriptome and the genome at the same time.

References:
1. Bray NL, et al. Nature Biotechnology. 2016 34(5):525–7.
2. Untergasser A, et al. Nucleic Acids Research. 2012;40(15):e115–e115.
3. Rice P, et al. Trends in Genetics. 2000(6):276–7.
4. Ye J, et al. BMC Bioinformatics. 2012;13(1).

Info

-	
filename	-
Figure	
-	
Availability	-

Corresponding Author	
Name, Surname	Ermes, Filomena
Email	ermes.filomena@gmail.com
Submitted on	02.05.2023