Clinical Immunology 180 (2017) 33–44

Contents lists available at ScienceDirect

Clinical Immunology

journal homepage: www.elsevier.com/locate/yclim

Systematic analysis of splicing defects in selected primary immunodeficiencies-related genes

Lucie Grodecká [a,b], Pavla Hujová [b], Michal Kramárek [a], Tereza Kršjaková [b], Tatiana Kováčová [a], Katarína Vondrášková [a], Barbora Ravčuková [b], Kristýna Hrnčířová [a], Přemysl Souček [a], Tomáš Freiberger [a,b,c,*]

[a] Central European Institute of Technology, Masaryk University, Brno, Czech Republic
[b] Molecular Genetics Laboratory, Centre for Cardiovascular Surgery and Transplantation, Brno, Czech Republic
[c] Institute of Clinical Immunology and Allergology, St. Anne's University Hospital and Masaryk University, Brno, Czech Republic

ARTICLE INFO

Article history:
Received 18 October 2016
Received in revised form 3 February 2017
accepted with revision 23 March 2017
Available online 27 March 2017

Keywords:
Splicing prediction
Splicing-affecting variants
Cryptic splice sites
Primary immunodeficiencies

ABSTRACT

Both variants affecting splice sites and those in splicing regulatory elements (SREs) can impair pre-mRNA splicing, eventually leading to severe diseases. Despite the availability of many prediction tools, prognosis of splicing affection is not trivial, especially when SREs are involved. Here, we present data on 92 *in silico*-/55 minigene-analysed variants detected in genes responsible for the primary immunodeficiencies development (namely *BTK*, *CD40LG*, *IL2RG*, *SERPING1*, *STAT3*, and *WAS*). Of 20 splicing-affecting variants, 16 affected splice site while 4 disrupted potential SRE. The presence or absence of splicing defects was confirmed in 30 of 32 blood-derived patients' RNAs. Testing prediction tools performance, splice site disruptions and creations were reliably predicted in contrast to SRE-affecting variants for which just ESRseq, ΔHZ_{EI}-scores and EX-SKIP predictions showed promising results. Next, we found an interesting pattern in cryptic splice site predictions. These results might help PID-diagnosticians and geneticists cope with potential splicing-affecting variants.

1. Introduction

Primary immunodeficiencies (PIDs) form a heterogeneous group of rare inborn diseases caused by improper function of a part of the immune system. Scientists currently recognize >260 different conditions, generally characterized by recurrent infections, inflammation and autoimmunity [1]. As the disease phenotypes may be similar for some conditions, it is often crucial for correct diagnosis and the most appropriate treatment to establish the particular underlying genetic defect. The detection of new genetic variants has been greatly facilitated by the introduction of high-throughput sequencing technologies. At the same time, these methods produce an immense number of sequence variants whose functional effect on gene expression (RNA processing and/or protein function) is unclear. These so called 'variants of unknown significance' (VUS) present a challenging task for all clinical geneticists.

As became increasingly clear during recent years, a significant portion of VUS affects the process of pre-mRNA splicing. Both the variants affecting conserved splice sites and those affecting much less conserved auxiliary splicing regulatory elements (SRE), *i.e.* splicing enhancers and silencers, often induce splicing aberrations. The former, however, are much easier to identify as potential splicing spoilers [2].

In parallel, performance of existing *in silico* prediction tools reflects the evolutionary conservation of the predicted elements: reasonably reliable predictions are given by the instruments assessing splice site changes while SRE-predicting tools results are generally less reliable and often hard to interpret [3–6]. However, no prediction tool has been found efficient so far in predicting an effect of splicing aberration, *i.e.* exon skipping, activation of a cryptic splice site (a consensus-like sequence that is not used as a splice site unless the normal process of splicing is disrupted by a mutation), intron retention, *etc.* [3]. Therefore, some splicing predictors can be useful for selecting potential splicing-affecting variants, but the diagnostic decisions still have to be based on *in vitro* analyses that reveal the actual effect of the variant at the level of cDNA [7]. *In vitro* testing should ideally be done on patient's RNA isolated from the affected tissue. Since this is often complicated by inaccessibility, other methods such as minigene splicing assays are frequently used [6]. So far, minigenes have helped to elucidate a number of splicing aberrations, yet their artificial nature still prompts the diagnosticians to cautiousness when working with these results [6].

In this report, we describe a systematic analysis of splicing effects of 92 sequence variants found in several PID-related genes. We found a good concordance between the results from patients' blood samples

Abbreviations: ESE, exon splicing enhancer; ESS, exon splicing silencer; NMD, nonsense-mediated decay; SRE, splicing regulatory element; VUS, variants of unknown significance.

* Corresponding author at: Centre for Cardiovascular Surgery and Transplantation, Molecular Genetics Laboratory, Výstavní 17/19, Brno 603 00, Czech Republic.
E-mail address: tomas.freiberger@cktch.cz (T. Freiberger).

http://dx.doi.org/10.1016/j.clim.2017.03.010
1521-6616/

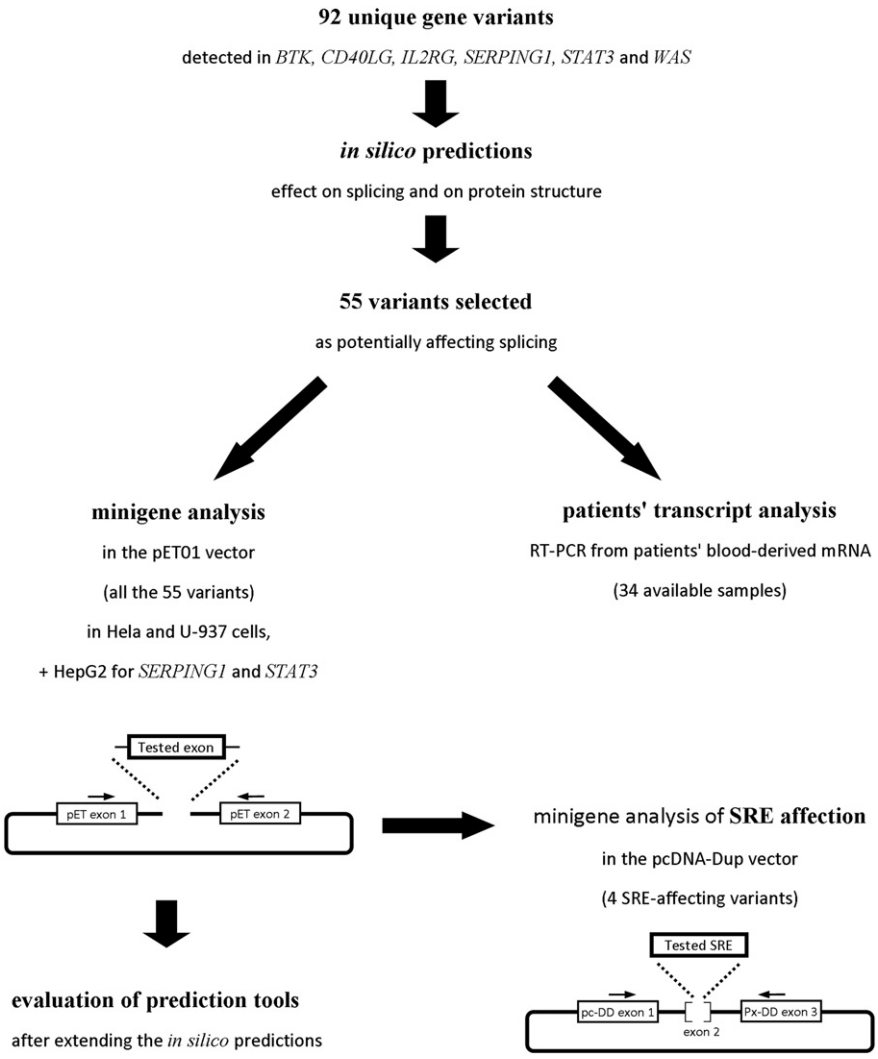

Fig. 1. Outline of the experimental procedures applied to the variants splicing analysis. The pET-derived minigenes are designated to display the effect of sequence variants on splicing of the respective exons and possibly its neighbouring introns. The pcDNA-Dup derived minigenes are intended to demonstrate SRE affection: inclusion of the middle exon into the mRNA is dependent on the presence of an ESE in the cloned sequence stretch. Abbreviations: pET: pET01 vector; pcDD: pcDNA-Dup vector.

and minigenes, mainly in assessing the existence of a splicing defect. Using *in silico* prediction tools, we show that frequently used cut-off for assessing predicted splice-site affecting variants (10% difference between wild type and mutant score) could detect vast majority of this type splicing aberrations. Further, we show the surprising efficiency of several prediction tools in selecting potentially SRE-affecting variants and for cryptic splice site predictions.

2. Materials and methods

2.1. Patients and nomenclature

The 92 sequence variants in the genes *BTK*, *CD40LG*, *IL2RG*, *SERPING1*, *STAT3*, and *WAS* were detected in patients undergoing mutation analysis for a suspected PID. The DNA variants' numbering is based on the transcripts depicted in Table S1. The nomenclature system follows the recommendations of the Human Genome Variation Society. The patients provided a written statement of informed consent. The research project was approved by the Ethics Committee of the Centre for Cardiovascular Surgery and Transplantation Brno.

2.2. Blood samples, RNA isolation and RT-PCR analysis on blood RNA

Patient RNA was isolated from peripheral blood samples stabilized with RNAlater (Thermo Fisher Scientific) using RiboPure-Blood Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. RT-PCR was carried out with SuperScript® One-Step RT-PCR System with Platinum® Taq DNA Polymerase (Thermo Fisher Scientific) for the genes *BTK*, *CD40LG* and *STAT3*. For the *SERPING1* and *WAS* genes, two step reactions were performed: RTs were carried out using Transcriptor First Strand cDNA Synthesis Kit with random hexanucleotides (Roche Applied Science) and consecutive PCRs were performed either with HotStartTaq Master Mix (Qiagen) or with Taq DNA polymerase (Qiagen). Primer sequences are depicted in Table S2 and specific conditions are available upon request. The resulting amplicons were visualized on 2% agarose gels and sequenced on ABI

Fig. 2. Splicing analysis of the variants shown to disrupt or create splice sites in pET minigenes. Results of minigene analysis in the pET01 vector and blood-derived patients' RNA analysis are shown as RT-PCR amplicons visualized on an agarose gel, next to the respective amplification schema. In the schemas, horizontal arrows depict primers position and bold vertical arrows depict approximate location of the analysed variants. The depicted minigene results were observed on RNA isolated from transfected HeLa cells. They are representative of at least 2 independent experiments. The identity of individual bands is described below the schemas and in Table 1. Note that some of the transcripts were detected only after TA-cloning and/or capillary electrophoresis, so they may not be visible on the gel. A) Splice-site affecting variants in the *BTK* gene. B) Splice-site affecting variants in the *IL2RG* gene. C) Splice-site affecting variants in the *SERPING1* gene. The asterisk denotes a band of heteroduplex sequence. D) Splice-site affecting variants in the *WAS* gene.