Clinical Immunology 160 (2015) 301–314

Contents lists available at ScienceDirect

Clinical Immunology

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

Application of whole genome and RNA sequencing to investigate the genomic landscape of common variable immunodeficiency disorders

Pauline A. van Schouwenburg [a,b,1], Emma E. Davenport [a,c,1], Anne-Kathrin Kienzler [a,b], Ishita Marwah [a,b], Benjamin Wright [a,c], Mary Lucas [a], Tomas Malinauskas [d], Hilary C. Martin [a,c], WGS500 Consortium [a,c,2], Helen E. Lockstone [c], Jean-Baptiste Cazier [c,e], Helen M. Chapel [a,b], Julian C. Knight [a,c,*,3], Smita Y. Patel [a,b,**,3]

[a] Nuffield Department of Medicine, University of Oxford, Oxford, UK
[b] Oxford NIHR Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK
[c] Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[d] Cold Spring Harbor Laboratory, W. M. Keck Structural Biology Laboratory, Cold Spring Harbor, NY 11724, USA
[e] Centre for Computational Biology, University of Birmingham, Haworth Building, B15 2TT Edgbaston, UK

ARTICLE INFO

Article history:
Received 25 April 2015
accepted with revision 20 May 2015
Available online 26 June 2015

Keywords:
Common variable immunodeficiency
Whole genome sequencing
Transcriptome
Polygenic
B-cell

ABSTRACT

Common Variable Immunodeficiency Disorders (CVIDs) are the most prevalent cause of primary antibody failure. CVIDs are highly variable and a genetic causes have been identified in <5% of patients. Here, we performed whole genome sequencing (WGS) of 34 CVID patients (94% sporadic) and combined them with transcriptomic profiling (RNA-sequencing of B cells) from three patients and three healthy controls. We identified variants in CVID disease genes *TNFRSF13B*, *TNFRSF13C*, *LRBA* and *NLRP12* and enrichment of variants in known and novel disease pathways. The pathways identified include B-cell receptor signalling, non-homologous end-joining, regulation of apoptosis, T cell regulation and ICOS signalling. Our data confirm the polygenic nature of CVID and suggest individual-specific aetiologies in many cases. Together our data show that WGS in combination with RNA-sequencing allows for a better understanding of CVIDs and the identification of novel disease associated pathways.

1. Introduction

Common Variable Immunodeficiency Disorders (CVIDs) are the most clinically prevalent primary antibody deficiencies, present in about 1 in 25,000 people [1]. CVID is characterised by recurrent infections, low serum levels of IgG, low IgA and/or IgM and poor specific antibody responses [2]. CVID is currently a diagnosis of exclusion leaving a highly heterogeneous patient group in terms of clinical features and complications. Some patients have a relatively mild phenotype, comprising of recurrent bacterial infections and infection-related complications while others suffer from disease related complications such as autoimmune cytopenias, polyclonal lymphoproliferation, enteropathy and lymphoid malignancy, indicating underlying immune dysregulation [3].

The heterogeneous disease phenotype of CVID has made resolving genetic aetiologies challenging. Causative variants of CVID-like primary antibody failure have been described in *CD19* [4–6], *CD21* [7], *CD81* [8], *CD20* [9], *ICOS* [10,11] and *TNFRSF13C* [12], conditions now classified as specific deficiencies in these genes (Table S1). Mutations in *CD27* [13,14], *PLCG2* [15,16], *LRBA* [17,18], *NFKB2* [19,20], *PIK3CD* [21,22] and *NLRP12* [23] cause CVID-like symptoms often combined with a more extensive clinical phenotype (Table S1). Variants in *TNFRSF13B* [24–26], *TNFRSF13C* [27], *FCGR2A* [28] and HLA [29] have been described to predispose to CVID (Table S1). Together these variants only explain the genetic cause of CVID-like diseases in very few patients and all genes were identified in familial cases of CVID, while the vast majority of CVID patients are sporadic. The wide variety in genes implicated in CVID further underlines the heterogenic nature of the disease. Further unravelling of the

Abbreviations: BCR, (B-cell receptor); CVIDs, (Common Variable Immunodeficiency Disorders); ESID, (European Society for Immunodeficiencies); GWAS, (Genome-wide Association Study); HLA, (Human Leucocyte Antigen); IPA, (Ingenuity Pathway Analysis); IVA, (Ingenuity Variant Analysis); LIR, (Leucocyte Ig receptor); PDB, (Protein Data Bank); PIDs, (Primary Immunodeficiency Disorders); NHEJ, (Non-homologous End Joining); RNA-seq, (RNA sequencing); SNVs, (single nucleotide variants); WGS, (whole genome sequencing); XLA, (X-linked agammaglobulinaemia).

* Correspondence to: J. C. Knight, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
** Correspondence to: S. Y. Patel, Oxford NIHR Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK.
E-mail addresses: julian@well.ox.ac.uk (J.C. Knight), smita.patel@ndm.ox.ac.uk (S.Y. Patel).

[1] The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors.
[2] A full list of members is provided in the Supplementary Material.
[3] The authors wish it to be known that, in their opinion, the last 2 authors should be regarded as joint Senior Authors.

http://dx.doi.org/10.1016/j.clim.2015.05.020

underlying genetic causes of sporadic CVID would give additional insight into the disease, opportunities for better patient stratification and novel insights into treatment opportunities.

In 2011 Orange et al. published the first genome-wide association study (GWAS) of CVID to identify genomic regions associated with CVID development [30]. Analysis of 363 patients and 3031 controls led to the conclusion that CVID is likely to be a polygenic disease with multiple novel susceptibility loci implicated. However, as of yet this has not resulted in further identification and elucidation of genes or variants that cause or predispose for sporadic CVID emphasizing the difficulties in studying this highly variable disease.

The development of next generation sequencing techniques has transformed the identification of the genetic basis of Mendelian diseases. In contrast, identification of the genetic basis remains challenging in polygenic conditions. Here, we present the first whole genome sequencing (WGS) data for a cohort of CVID patients to investigate novel underlying aetiologies. We further leveraged the potential of WGS by combining the results with global transcriptomic profiling through RNA-sequencing (RNA-seq). Because of the complex and probable polygenic nature of CVID, we combine the identification of genes of interest with pathway-based analysis and focus on combining these results to identify pathways dysregulated in CVID.

2. Material and methods

2.1. Samples

Patients were recruited into the study through the Clinical Immunology Department at the Oxford University Hospital, Oxford. All patients gave informed written consent and the studies were performed according to the Declaration of Helsinki. All 34 patients were of Caucasian origin and met the ESID diagnostic criteria at the time of enrollment [2]. The majority of patients were regularly followed in the Clinical Immunology clinic at 6 monthly intervals over a period of up to 30 years with detailed clinical information entered into the local database that enabled accurate clinical phenotyping. A summary of the clinical phenotype and laboratory characteristics of the patient cohort can be found in Table 1 and a more complete overview can be found in Table S2.

2.2. WGS500

A cohort of 34 CVID patients was selected for WGS as part of the WGS500 project [31]. This is a collaborative project between the University of Oxford and Illumina, which aims to sequence the genomes of 500 individuals with a range of diseases including rare inherited diseases,

Table 1
Overview of clinical information on the 34 CVID patients.

Patient no	Sex	Age at onset of symptoms	CD19 count[a]	CD3 count[a]	CD4 count[a]	CD8 count[a]	IgG levels[a]	IgM levels[a]	IgA levels[a]	Clinical phenotype[3]
Normal			0.1–0.5	0.7–2.1	0.3–1.4	0.2–0.9	6.0–16.0	0.4–2.5	0.8–3.0	
C003	F	54	↓↓	↓	Normal	Normal	↓↓	Normal	↓↓	Polyclonal lymphoproliferation
C006	M	34	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	Polyclonal lymphoproliferation
C018	M	8	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
C019	M	18	Normal	Normal	Normal	↑	↓↓	↓↓	↓↓	No disease-related complications
C023	M	25	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
C024	F	48	Normal	↓	Normal	↓	↓↓	↓↓	↓↓	No disease-related complications
C028	F	47	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	Polyclonal lymphoproliferation
C032	M	26	↓	Normal	↓	↓	↓	Normal	↓↓	Polyclonal lymphoproliferation
C036	F	19	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
C038	M	5	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
C044	F	31	Normal	Normal	Normal	Normal	↓	↓↓	↓	Polyclonal lymphoproliferation; autoimmune cytopenias
C063	F	48	Normal	Normal	Normal	Normal	↓↓	Normal	↓↓	No disease-related complications
C065	M	5	Normal	Normal	Normal	Normal	↓	↓	↓↓	No disease-related complications
C072	F	34	↓↓	Normal	Normal	Normal	↓↓	↓/↓↓	↓↓	No disease-related complications
C078	M	13	Normal	Normal	Normal	Normal	↓↓	Normal	↓↓	No disease-related complications
C082	F	11	↓↓	Normal	Normal	Normal	↓↓	↓↓	↓↓[b]	Autoimmune cytopenias
C085	F	37	Normal	Normal	Normal	Normal	↓	↓↓	↓↓	No disease-related complications
C089	F	17	↓↓	Normal	Normal	Normal	↓↓	↓	↓↓	No disease-related complications
D038	M	22	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
D209	M	41	Normal	Normal	Normal	Normal	↓↓	↓	↓↓	No disease-related complications
D232	F	44	↑	↑	↑	Normal	↓	Normal	↓↓	No disease-related complications
D269	F	44	Normal	↓	↓	Normal	↓↓	Normal	↓↓	Polyclonal lymphoproliferation; autoimmune cytopenias
D276	M	1	↓↓↓	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
D325	F	35	Normal	↓	Normal	Normal	↓↓	↓	↓↓	Polyclonal lymphoproliferation; autoimmune cytopenias
D334	M	15	Normal	Normal	Normal	Normal	↓↓	↓	↓↓	No disease-related complications
D345	F	20	Normal	↑	Normal	↑	↓↓	↓	↓↓	No disease-related complications
D374	M	3	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
D575	M	2	Normal	Normal	Normal	Normal	↓↓	↓↓	↓↓	No disease-related complications
D641	F	20	Normal	↓	↓	Normal	↓↓	↓/↓↓	↓↓	No disease-related complications
D667	M	13	Normal	Normal	Normal	Normal	↓↓	↓	↓↓	No disease-related complications
D705	F	49	Normal	Normal	Normal	Normal	↓/↓↓	↓/↓↓	Normal/↓	No disease-related complications
D745	M	5	Normal	↑	↑	↑	↓↓	Normal	↓↓	No disease-related complications
D765	F	30	Normal	Normal	Normal	Normal	↓↓	Normal	↓↓	No disease-related complications
D839	M	25	Normal	↓	Normal	Normal	↓↓	↓	↓↓	Polyclonal lymphoproliferation

↓ = between lower normal and half lower normal values; ↓↓ = less than half lower normal value; ↓↓↓ = undetectable; ↑ = between upper normal value and twice upper normal value;↑↑ = more than twice upper normal value.

[a] Values at diagnosis.

[b] Values at diagnosis not available; values listed are post-infusion of intravenous immunoglobulin, which does not alter IgA levels.