

Clinical Research

Genetic diagnosis of autoinflammatory disease patients using clinical exome sequencing

Laura Batlle-Masó^{1,2}, Anna Mensa-Vilaró^{3,4}, Manuel Solís-Moruno^{1,2}, Tomàs Marquès-Bonet^{1,5,6,7}, Juan I. Arostegui^{3,4,8}, Ferran Casals²

¹ Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Doctor Aiguader 88, Barcelona, Spain.

² Genomics Core Facility, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Parc de Recerca Biomèdica de Barcelona, 08003 Barcelona, Spain.

³ Department of Immunology, Hospital Clínic, Barcelona, Spain.

⁴ Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.

⁵ Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010, Barcelona, Spain.

⁶ CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldori i Reixac 4, 08028 Barcelona, Spain.

⁷ Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain.

⁸ Universitat de Barcelona, Barcelona, Spain

Address for correspondence:

Ferran Casals

Genomics Core Facility, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra

Parc de Recerca Biomèdica de Barcelona, 08003 Barcelona, Spain.

ferran.casals@upf.edu

+34 933 160 846

Abstract

Autoinflammatory diseases comprise a wide range of syndromes caused by dysregulation of the innate immune response. They are difficult to diagnose due to their phenotypic heterogeneity and variable expressivity. Thus, the genetic origin of the disease remains undetermined for an important proportion of patients. Targeted sequencing of candidate genes can be a cost-effective approach to identify likely causative genetic variants, which will contribute to the diagnosis and better understanding of the disease.

We aim to identify causal genetic variants in patients with autoinflammatory disease and to test the advantages and limitations of the clinical exome gene panels for molecular diagnosis. Twenty-two unrelated patients with clinical features of autoinflammatory diseases were analysed using clinical exome sequencing (~4,800 genes), followed by bioinformatic analyses to detect likely pathogenic variants.

By integrating genetic and clinical information, we found a likely causative genetic variant in *NFKBIA* (p.D31N) in a North-African patient with deficit of interleukin-1 receptor antagonist, and *DNASE2* (p.G322D) in a Spanish patient with a suspected lupus-like monogenic disorder. Also, we found variants likely to increase the susceptibility to autoinflammatory diseases in three additional Spanish patients: one with an initial diagnosis of juvenile idiopathic arthritis who carries two *UNC13D* variants (p.R727Q and p.A59T), and two with early-onset inflammatory bowel disease harbouring *NOD2* variants (p.L221R and p.A728V respectively). Our results show a similar proportion of molecular diagnosis to other studies using whole-exome or targeted resequencing in primary immunodeficiencies. Thus, despite its main limitation of not including all candidate genes, clinical exome targeted sequencing can be an appropriate approach to detect likely causative variants in autoinflammatory diseases.

Keywords

Autoinflammatory diseases, primary immunodeficiency, targeted sequencing, clinical exome, gene panels.

INTRODUCTION

Autoinflammatory diseases (AID) constitute a family of immune disorders characterized by recurrent episodes of sterile inflammation that may appear without a known trigger. AID are considered primary immunodeficiency diseases (PID), a wide classification which includes different subgroups, all of them caused by dysregulations of specific pathways of the immune system. Although reliable epidemiological data is only available for some of the subgroups (i.e. inflammosomopathies), all AID are considered rare diseases (Wekell et al., 2017). Clinically, they are characterized by recurrent episodes of fever, systemic inflammation and other symptoms such as skin rashes, abdominal pain, chest pain, lymphadenopathy or arthritis (Wekell et al., 2016). The severity of the disease is variable, supporting the observations of clinical heterogeneity that makes a definitive diagnosis difficult, especially in early-onset patients without a positive family history (Almeida de Jesus and Goldbach-Mansky, 2013).

Over the past decades, several efforts were focused on better definition and classification of the wide variety of phenotypes and clinical manifestations included in AID group (Ciccarelli et al., 2014; Wekell et al., 2016). Many molecular mechanisms underlying these diseases have been discovered, most of them affecting genes related to the inflammasome (*TNFRSF1A*, *NLRP3* and *MEFV* among others (Hoffman and Broderick, 2016)). All these advances represented a significative improvement in diagnosis rates. However, there is still a significant proportion of patients with clinical features suggestive of AID that remain without a clear molecular diagnosis (Martorana et al., 2017; Notarangelo and Sorensen, 2008).

Current massive sequencing technologies have raised the possibility of analysing the whole genome or whole exome of patients. Alternatively, in targeted sequencing studies the analysis is restricted to a set of previously defined candidate genes, thus reducing economic costs and data analysis efforts (Seleman et al., 2017). In recent years, targeted sequencing panels have become a useful tool to identify the molecular causes of autoinflammatory diseases (Karacan et al., 2019; Omoyinmi et al., 2017; Papa et al., 2019; Yao et al., 2016) and other primary immunodeficiencies (reviewed in Yska et al., 2019).

In this project we assess the suitability and performance of a gene panel approach, the TruSight One Sequencing Panel (Illumina), to detect the genetic origin in presumably monogenic AID. We first analysed the proportion of candidate genes for autoinflammatory disorders included and properly sequenced in the panel. We discussed their applicability and its benefits or limitations compared to other approaches, such as whole genome or whole exome sequencing. Second, we performed targeted sequencing in 22 patients, and exhaustively analysed the data to detect potentially causative variants, considering different genetic models for the disease and the available clinical information for each patient.

MATERIALS AND METHODS

Study cohort

This study includes 22 unrelated patients (named A1-A22) with severe symptoms and clinical suspicion of AID. Written informed consent to participate in the study in accordance with the Declaration of Helsinki was obtained before the extraction of blood samples. The protocol was approved by the Comitè Ètic d'Investigació Clínica-Parc de Salut Mar (Barcelona). Genomic DNA was extracted from whole blood using AllPrep DNA/RNA kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. DNA

quality and concentrations were measured using Nanodrop spectrophotometer (Thermo Fischer Scientific, Waltham, Massachusetts, USA), and Qubit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) or PicoGreen (Invitrogen, Carlsbad, California, USA) fluorimetric quantification methods.

DNA library preparation and sequencing

Targeted sequencing capture was performed using the TruSight One Sequencing Panel (Illumina, San Diego, California, USA) which includes 4,813 genes related to disease. After the enrichment, quality control of the pooled library was performed with the Bioanalyzer System (Agilent, Santa Clara, California, USA) and quantified with a specific qPCR for library adaptors (Kappa Biosystems-Roche, Basel, Switzerland). Sequencing was performed on a NextSeq 500 System (Illumina, San Diego, California, USA) in a 2 × 75 paired-end cycles high-output run.

Data analysis

Bcl output files were converted to Fastq and demultiplexed with bcl2fastq (Illumina, San Diego, California, USA). Mapping and alignment against the human reference genome (GRCg38/hg38) was performed with BWA-MEM. Picard (<http://www.picard.sourceforge.net>) was used to reorder and sort SAM files before marking PCR duplicates and replace read groups. Following the Genome Analysis Toolkit (GATK (McKenna et al., 2010)) best practices we performed base quality score recalibration and local indel realignment before the variant calling. The resulting VCF file was annotated using Anovar (Wang et al., 2010).

We used VarScan2 (Koboldt et al., 2012) to analyze somatic variation and XHMM (Fromer et al., 2012; Fromer and Purcell, 2014) for copy number alterations. The study of synonymous variants was performed using the Transcript-Inferred Pathogenic Score from <http://trap-score.org/about.jsp> (TraP(Gelfman et al., 2017)).

Genetic variant filtering and prioritizing

A first quality filter discarded variants with $DP < 5$ and $QUAL < 30$ (depth and quality given by GATK). We also excluded variants present in more than ten patients to avoid possible sequencing artifacts. We removed indels up to 10 base pairs next to another indel due to its high probability of being produced by mapping errors. Intronic variants not involved in splicing procedures were also discarded. We used the bed file provided by the manufacturer to exclude off-target regions and a public dataset to detect genes which are known to be enriched in sequencing errors (Fuentes Fajardo et al., 2012).

As shown in **Figure 1** we applied different filters considering the different groups of variants and gene properties. For loss of function variants, we applied a filter of $MAF < 1\%$ in gnomAD (Lek et al., 2016) database. For these variants, we also considered the constraint metrics of gnomAD such as the precalculated probability of being loss-of-function intolerant (pLI) for each gene. We retrieved information about the function of the gene and its implication in human disease using the Gene Ontology (GO (Ashburner et al., 2000; The Gene Ontology Consortium, 2019)). Missense variants were filtered by allele frequency ($MAF < 1\%$) and GO terms, as well as conservation and deleteriousness scores such as GERP (Chong et al., 2015; Cooper et al., 2005) and CADD (Kircher et al., 2014), retaining variants with values higher than 2 and 10, respectively. Genes with missense variants were additionally annotated with their haploinsufficiency scores (Huang et al., 2010; Steinberg et al., 2015), intolerance to genetic variation (RVIS (Petrovski et al., 2015)) and essentiality scores (Khurana et al., 2013) to detect single heterozygous variants which caused a gain of function effect. This mechanism has been widely described in AID and it has been related to the variable clinical expression of the phenotype (Martorana et al., 2017). Finally, other *in silico* predictors of the protein effect were also used as a proxy of pathogenicity (PolyPhen2 (Adzhubei et al., 2010), Provean (Choi et al., 2012) and SIFT (Kumar et al., 2009)). Synonymous variants with $MAF < 0.1\%$ were considered as potential candidate variants. We annotated it using TraP score

(Gelfman et al., 2017) and filtered according to the recommended threshold (Benign if $\text{TraP} < 0.446$) and $\text{MAF} < 0.1\%$.

Copy number variation was studied using XHMM (Fromer et al., 2012; Fromer and Purcell, 2014) and filtered afterwards according to the best practices described by the developers. We also used the dataset provided by Audano et al (Audano et al., 2019) to filter out variants described at high frequencies in the population. Finally, possible somatic variants were called using VarScan2 (Koboldt et al., 2012) and filtered considering its variant allele frequency (VAF) and overall quality of the calling (p-value given by VarScan2).

As candidate genes for AID and PID, we used those in the list provided by the International Union of Immunology Societies (IUIS) (Picard et al., 2018). For each patient, we performed an extensive clinical review of the phenotype and family history. Finally, candidate variants were visually inspected in the respective alignment using Integrative Genome Viewer (IGV (Robinson et al., 2011)). We filtered out variants only supported with poor quality reads or low mapping quality.

Validation of somatic variants

We performed ADS (Becker et al., 1996) to validate the presence of somatic variation and to calculate their frequencies. We included two positive controls to validate the precision of the technique (CP1 and CP2), both cases have been previously reported (Mensa-Vilaró et al., 2019).

Nine independent amplicons using three different sets of primers were obtained for each variant using PCR. Sequencing of amplicons was performed on a MiSeq equipment (Illumina, San Diego, California, USA) in a 2×250 paired-end cycles Nano flow cell run. Fastq files were mapped to the human reference genome (GRCh38/hg38) using BWA and each of the resulting BAM files was inspected using Pysam Python's package (Li et

al., 2009). Statistical analysis was performed with RStudio (RStudio Team, 2016) in order to distinguish sequencing errors from true somatic variation.

RESULTS

AID-associated genes in the TruSight One Sequencing Panel

The TruSight One Sequencing Panel, also referred to as clinical exome gene panel, consists of 4,813 genes associated with human disease. With the aim of testing its applicability and performance to study AID, we analyzed the ability to capture and produce reliable sequencing data for candidate genes. We first assessed the proportion of genes related to autoinflammatory disorders included in this panel. For that, we used 35 genes described to cause AID by the IUIS (Picard et al., 2018) plus 31 additional genes selected from the literature, for a total of 66 genes (Autoinflammatory candidate genes, Acg). We also analyzed 298 genes associated with other immune diseases by the IUIS (Other PID candidate genes, OPcg). 68% of the Acg (45/66) and 77% of the OPcg (231/298) are included in the TruSight One Panel (**Figure 2, Table S1**).

The mean coverage per sample was 108x, with a maximum of 149x (sample A13) and a minimum of 70x (sample A10). We performed two different analyses in order to assess the sequencing performance for the genes included in the panel. First, we analyzed the depth of coverage (average number of reads for targeted positions). For this, we measured the mean coverage per gene (targeted sequence) and we classified it into three categories: high depth (>50x), medium depth (20-50x) and low depth (<20x) (**Table S2**). 95.56%, 95.67% and 95.88% of the genes were sequenced at high depth for Acg, OPcg, and the rest of the genes of the panel, respectively (**Figure 2**). Two genes were sequenced at low depth among the Acg (*PSMB8*, *PSMB9*) and three on the OPcg list (*CFB*, *TAP1* and *TAP2*).

Next, we analyzed the horizontal coverage. We calculated the mean number of base pairs uncovered in each gene for at least 80% of the patients and we classified it as complete (all bases sequenced), good (1-50 unsequenced bases) and poor (>50 unsequenced bases) (**Table S3**). The horizontal coverage was complete or good for most of the Acg (95.56%, 43/45), as well as for OPcg (98.26%, 227/231). The genes with more than 50 unsequenced bases were *P2RX7* and *SH3BP2* for Acg, and *C4B*, *CFB*, *TAP1* and *TBX1* for OPcg. For the rest of the panel, 95.57% of the genes were complete (4,337 out of 4,538) (**Figure 2**).

Genetic variants in patients with suspected AID

We explored different possible scenarios to identify potential causal genetic variants in the 22 enrolled patients. For a dominant model, we first analyzed the presence of loss of function variants which are, a priori predicted to have an important effect on the protein function and are therefore good candidates to be at the origin of genetic diseases. We included in this group nonsense variants (stop-codon), frameshift indels, and splice site disrupting variants (MacArthur et al., 2012; MacArthur and Tyler-Smith, 2010). We filtered out the variants with MAF>1% in gnomAD (Lek et al., 2016) and we next analyzed the variants considering the gnomAD precomputed probability for genes of being loss-of-function intolerant (pLI). We also considered the GO terms description and kept those variants in genes related to the immune system or the inflammatory response. After this, we generated a list of 18 candidate LoF variants for AGcg and OPcg, all of them heterozygous (**Table S4**). Any of the variants was considered as a potential candidate after the clinical review of the patient's phenotype.

We expanded the analysis to missense genetic variants, with a list of 116 mutations with MAF<1% (**Table S5**). We used deleteriousness (CADD>10) and conservation values (GERP>2) to filter missense variants. We also considered variant features such as *in silico* analyses to prioritize all detected variants, and the inheritance pattern of each

gene described by IUIS (Picard et al., 2018) and OMIM (<http://www.omim.org/>). We used essentiality and haploinsufficiency scores to better detect single heterozygous variants causing a gain of function effect. We propose two candidate heterozygous missense variants in *NFKBIA* (patient A11) and *DNASE2* (patient A21). Both variants are novel and show high CADD and GERP scores and detrimental effects were predicted by most *in silico* tools (**Table 1**).

We next considered an autosomal recessive model including both homozygous and compound heterozygous genotypes. No candidate homozygous variant was in any candidate gene. For the compound heterozygotes analysis, we considered together missense and other functional variants. Patient A10 has two variants located in *UNC13D*, a gene associated with hemophagocytic lymphohistiocytosis. Sieni et al (Sieni et al., 2011) described a patient with these same variants and suggested that they could be considered as susceptibility factors. We also described several variants located at *NOD2* in patients A19 and A20. This gene is associated with autoinflammatory disorders and it is well known to play an important role in this kind of disease (Negroni et al., 2018). A19 and A20 show a clinical picture concordant with alterations in this gene (**Table 2**). Based on pathogenicity scores and literature search, most of the *NOD2* variants described in patients A19 and A20 were considered to be most probably benign (Hugot et al., 2001), except the two missense variants shown in **Table 1**.

We also analyzed the presence of synonymous changes in candidate genes. Synonymous variants are traditionally excluded from disease genetics analysis, and are discarded when filtering for functional variants. However, synonymous variants have been reported to be related to disease in some cases, especially if they provoke the appearance of novel, cryptic splicing sites (Simons et al., 2015). As an example, a novel synonymous mutation has been reported as causal for IL-7R deficiency because of its effect on splicing (Gallego-Bustos et al., 2016). We filtered variants using a very restrictive allele frequency of $MAF < 0.1\%$ or non-described at gnomAD. We used the

TraP Score (Gelfman et al., 2017) to assess its pathogenicity. All the values were below the established threshold of 0.446 for pathogenicity. The synonymous variant with the highest TraP score (0.389, probably benign) was found in *NFKBIA* in patient A11. As described above, an interesting missense variant was found in this same patient in the same gene.

Finally, we analyzed copy number variation usingXHMM software (Fromer et al., 2012; Fromer and Purcell, 2014). Massive parallel sequencing technologies are limited to detect structural variants due to the short length of reads. Therefore, many of our calls were discarded due to low-quality values. We removed CNV's present in more than 10% of the patients due to the lack of similarities in their clinical phenotypes. We also compared our candidates with a recent structural variation reference dataset (Audano et al., 2019) to discard the variants which have been already described at high frequencies in the population. After this, no candidate variants were in Acg and OPcg.

Somatic variation

Recent studies corroborated that gene mosaicism is an important mechanism in the pathogenesis of AID (Mensa-Vilaró et al., 2019). Thus, we considered the possibility of that the disease originated by a somatic variant. While germline variants are expected to be present in approximately 50% of the reads, somatic variants will typically depart from these expectations and will show lower variant allele frequencies (VAF). Because of the low number of reads supporting them, these variants are commonly discarded by variant calling software. With the aim of detecting possible somatic variants, we use VarScan2 which has shown good performance for these analyses and allows reduction in the allelic imbalance thresholds (Suzuki et al., 2017).

We generated a table with 12 variants located in Acg and OPcg (**Table S6**), two of them present in two of the patients. We performed ADS (see Methods) to validate these variants, and to calculate their VAF. We included two previously reported somatic

variants as positive controls (CP1 and CP2) (Mensa-Vilaró et al., 2019). We validated the presence of somatic variants in CP1 and CP2 with similar VAF as the previously described. We also compared the real VAFs to the ones given informatically by GATK Haplotype Caller and VarScan2 (**Figure S1, panel A**). The candidate somatic variants in enrolled patients presented VAFs between 1.25 and 9.42%. We normalized the values by comparing the frequencies of the alternative alleles to the other two non-reference alleles. For most of the samples, the alternative allele frequency was indistinguishable from the other non-reference alleles, suggesting a false-positive. Only in patient A7, who carries a missense *NLRP3* variant (p.G566D), we detected a slight increase in the number of reads supporting the alternative allele when compared to the other two non-reference alleles. Due to the concordance between the patient's phenotype and clinical pictures caused by *NLRP3* described alterations, we performed ADS of this variant in sample A7 together with the previous CP1 and CP2 samples plus two additional DNA samples from healthy individuals (HI1 and HI2). We merged the values of each triplicate (see Methods), and the two samples in each control group, classifying the results in three categories: A7, CP and HI. After normalization, we performed a t-test and calculated the p-values (**Figure S1, panel B**). The percentage of reads supporting the alternative allele was 0.069%, 0.073% and 0.082% for A7, CP and HI respectively. The difference between groups was not significant (p-value > 0.05). This indicates that the slight increase of alternative reads in sample A7 that we saw in the first batch of ADS experiments was not caused by the presence of a somatic variant.

DISCUSSION

We performed targeted sequencing of ~4,800 genes of medical interest in 22 samples from AID patients, mostly with early-onset. We propose a candidate gene variant in two of the patients, based on stringent filtering of DNA variants and the patients' clinical phenotypes. For three additional patients, we also propose candidate variants which can be acting as susceptibility factors of the disease.

Patient A11 was a Spanish male diagnosed at birth of deficiency of interleukin-1 receptor antagonist (DIRA), and he died soon after birth because of multiorgan failure. Our analysis revealed the patient carried a novel missense heterozygous variant at *NKFBIA* (p.D31N). *NKFBIA* is an inhibitor of the NF κ B family which acts as a regulatory element of the inflammatory response (Hayden and Ghosh, 2008). When proinflammatory stimuli are detected *NKFBIA* is phosphorylated and degraded, leading to the nuclear translocation of NF κ B (Ali et al., 2013). DiDonato et al (DiDonato et al., 1996) described that phosphorylation occurs at p.S32 and p.S36, however, they did not find evidence of a functional role for p.D31 (DiDonato et al., 1996). Mutations in *NKFBIA* are associated with ectodermal dysplasia in immunodeficiency (Courtois et al., 2003), which correlates with our patient phenotypes. The patient presented with sterile periostitis, pustular skin, rash, and soft tissue swelling. We suspect that the variant p.D31N could be contributing to the disease in patient A11. This variant fits to some of the standard criteria (Richards et al., 2015) to be classified as pathogenic: absence in databases; presence in a critical functional domain with reported pathogenic variants; in a gene where missense variation is a common mechanism of disease; a significant lack of missense and LoF variants; deleterious effect according to *in silico* predictors; and specific gene phenotype. Unfortunately, we were not able to perform segregation studies to validate the possible *de novo* origin due to the impossibility of obtaining new samples. We reported the variant to ClinVar (SCV001147010). Of interest, this patient also harbours a synonymous variant in homozygosity with a TraP score close to the pathogenicity threshold, a high CADD value of 17.04 and located in a promoter region (ENSR00000067656).

Patient A21 is a Spanish male with early-onset lupus-like symptomatology. A novel missense heterozygous variant at *DNASE2* was found in the patient (p.G322D), this variant may be pathogenic (Richards et al., 2015) due to its absence in databases, its *in silico* predictors, conservation values (**Table 1**), and reported pathogenic variants in the gene (Rodero et al., 2017). *DNASE2* has recently been associated to inflammatory

diseases (Rodero et al., 2017), and mutations in *DNASE1* are well known to be causing monogenic forms of systemic lupus erythematosus (Martínez Valle et al., 2008; Yasutomo et al., 2001), with a clinical phenotype partially overlapping with that described in this family. In our patient, the disease seems to follow an autosomal dominant pattern of inheritance, as some close relatives are also affected (father, uncle and grandmother). This diverges from the described autosomal recessive cases. We performed Sanger sequencing in the patient (**Figure S2**) and we propose the variant p.G322D as a good candidate to underly the disease in patient A21. We reported the variant to ClinVar (SCV001147011).

In addition to the analysis to detect disease-causing variants, we also considered the possibility of mutations contributing to the phenotype although not responsible for the whole clinical picture. Patient A10 is a 12 year-old girl from North Africa with systemic-onset juvenile idiopathic arthritis (JIA). She suffered episodes of macrophage activation syndrome (MAS), which are associated with *UNC13D* mutations. We found two variants at *UNC13D* which have been already reported by Sieni et al (Sieni et al., 2011) in the analysis of haemophagocytic lymphohistiocytosis. Both variants are heterozygous and missense (p.R727Q and p.A59T), with low population allele frequencies (0.05% and 1.6% respectively) and high predicted impact values (CADD>20). Although the patient's main clinical diagnosis is not macrophage activation syndrome (MAS), the patient's phenotype resembles the ones described in Sieni's work. This led us to suspect that maybe part of the phenotype is caused by these missense variants. Of note, the presence of two or more causal genes for different Mendelian disorders has been reported in 4.9% of the 2,076 molecularly diagnosed patients in an exhaustive analysis of more than 7,000 cases (Posey et al., 2016). The case reported here would fit the situation where the two disease phenotypes in the same individual are partially overlapping.

We also described many variants in *NOD2* gene in our cohort. *NOD2* has a key role in immune function and inflammatory pathways. Numerous variants have been described in the gene in recent years, mainly associated with Blau syndrome and Crohn disease (Negroni et al., 2018). Of interest, two of the patients with *NOD2* variants (A19 and A20) have early-onset inflammatory bowel disease, a phenotype which has already been associated to *NOD2* gene (Cho and Abraham, 2007; Negroni et al., 2018). Each of them carries one missense variant of uncertain significance, p.L221R on patient A19 and p.A728V on A20. Both variants have been reported by Hugot et al (Hugot et al., 2001) in inflammatory bowel disease patients accompanied by other coding variants. We hypothesize a possible additive effect which would contribute to the final phenotype.

Of interest, in this work, we show that beyond the detection of germline genetic variants, targeted sequencing also allows the detection of somatic variation if relatively high sequencing depths are achieved. Thus, while no important differences are seen in the success rate of whole-exome sequencing versus candidate genes approaches (see below), reducing the number of targeted genes allows increased depth at still reasonable costs, which is a critical aspect for somatic variant detection. It can be of special interest in the case of PID in general and AID in particular, where somatic variants will probably be present in the blood or even saliva, the two most common starting materials for DNA extractions. We recommend using software allowing the reduction of the thresholds on the number of reads supporting the alternative allele (i.e. VarScan2) followed by ADS to validate candidate variants. ADS experiments increase the coverage of the targeted region from the range of hundreds to thousands, allowing us to discard false-positives.

The proportion of patients with a proposed molecular diagnosis in this work is similar to other previous analyses in PID, both from WGS-WES (Arts et al., 2019; de Valles-Ibáñez et al., 2018) and targeted sequencing approaches (see Introduction). However, the number of diagnosed patients is barely comparable across studies because of the use

of different filters and stringency criteria to consider a genetic variant as causal. This would ultimately require functional validation experiments, which are mostly performed in a reduced subset of the candidate variants. Also, the proportion of monogenic cases in a cohort depends on other features, such as disease severity, the age of onset and the disease prevalence (de Valles-Ibáñez et al., 2018). Thus, the different inclusion criteria of patients, can also partially explain the differences in the success rates in detecting causal genes across studies. In addition to the clinical exome approach being limited to a particular gene set, other technical constraints refer to the non-inclusion of regulatory regions and limited power for structural variant detection. We note some of these factors are also common to whole-exome or even whole-genome strategies. Beyond technical and analytical reasons, we should always consider the possibility that the disease follows non-monogenic models, such as a digenic pattern of inheritance (Ameratunga et al., 2018; de Valles-Ibáñez et al., 2018; Hoyos-Bachiloglu et al., 2017) or gene mosaicism (Mensa-Vilaró et al., 2019).

Overall, we propose that targeted sequencing approaches of candidate genes, such as clinical exome, is a powerful tool for the genetic diagnoses of AID. This approach allows the analysis of regions of interest with high sequencing depths. Its limitation to specific regions is also seen in several whole exome and genome data studies, where in spite of generating more complete data, the analysis is often restricted to a set of candidate genes, at the cost of reducing sequencing depth (Meyts et al., 2016). In this sense, the design of specific gene panels for patients with similar disorders can represent an efficient strategy for diagnosis, which may even be able to capture the presence of somatic mosaicism which is of importance in diseases such as AID. Of importance, sequencing coverage and depth analyses to detect genes that have not been properly uncovered should be performed to avoid false-negative results. Finally, the creation of an adequate candidate genes list is critical for this approach and should be based on a

strong knowledge of the disease pathophysiology and the use of available tools to describe functionally related genes.

Acknowledgements

We thank all participants in the study. LBM is supported by a Formació de personal Investigador fellowship from Generalitat de Catalunya (2018_FI_B00072). MSM. is supported by the María de Maetzu Programme (MDM-2014-0370-16-3). This study was funded by grants SAF2012-35025 and SAF2015-68472-C2-2-R from the Ministerio de Economía y Competitividad (Spain) and FEDER (EU) to FC; RTI2018-096824-B-C22 grant from the Spanish Ministry of Science, Innovation and Universities co-financed by ERDF (FC), by Direcció General de Recerca, Generalitat de Catalunya (2014SGR-866 and 2017SGR-702) to FC, the CERCA Programme/Generalitat de Catalunya (JIA), SAF2015-68472-C2-1-R grant from the Spanish Ministry of Economy and Competitiveness co-financed by European Regional Development Fund (ERDF) (JIA), RTI2018-096824-B-C21 grant from the Spanish Ministry of Science, Innovation and Universities co-financed by ERDF (JIA), and AC15/00027 grant from the Instituto de Salud Carlos III / Transnational Research Projects on Rare Diseases (JIA).

Conflict of interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Figure Legends

Figure 1. Overview of the filtering process. HI, haploinsufficiency values. Italics indicate the used software.

Figure 2. Selection of candidate genes and presence at the TruSight One Sequencing Panel. Depth of coverage, calculated as the mean number of reads per gene. High: genes with mean >50; Medium: 20-50x; Low: <20x. Horizontal coverage, number of uncovered bases per gene in at least 80% of the patients. Complete: no uncovered bases; Good: 1-50 uncovered bases; Poor: >50 uncovered bases.

Table 1. Candidate genetic variants in patients with autoinflammatory disorders.

Table 2. Clinical information of patients harboring variants likely to be causing the autoinflammatory disorder.

REFERENCES

- Adzhubei, I. a, Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. <https://doi.org/10.1038/nmeth0410-248>
- Ali, S., Hirschfeld, A.F., Mayer, M.L., Fortuno, E.S., Corbett, N., Kaplan, M., Wang, S., Schneiderman, J., Fjell, C.D., Yan, J., Akhabir, L., Aminuddin, F., Marr, N., Lacaze-Masmonteil, T., Hegele, R.G., Becker, A., Chan-Yeung, M., Hancock, R.E.W., Kollmann, T.R., Daley, D., Sandford, A.J., Lavoie, P.M., Turvey, S.E., 2013. Functional genetic variation in NFKBIA and susceptibility to childhood asthma, bronchiolitis, and bronchopulmonary dysplasia. *J. Immunol.* 190, 3949–58. <https://doi.org/10.4049/jimmunol.1201015>
- Almeida de Jesus, A., Goldbach-Mansky, R., 2013. Monogenic autoinflammatory diseases: concept and clinical manifestations. *Clin. Immunol.* 147, 155–74. <https://doi.org/10.1016/j.clim.2013.03.016>
- Ameratunga, R., Woon, S.T., Bryant, V.L., Steele, R., Slade, C., Leung, E.Y., Lehnert, K., 2018. Clinical implications of digenic inheritance and epistasis in primary immunodeficiency disorders. *Front. Immunol.* 8. <https://doi.org/10.3389/fimmu.2017.01965>
- Arts, P., Simons, A., AlZahrani, M.S., Yilmaz, E., Alldrissi, E., van Aerde, K.J., Alenezi, N., AlGhamdi, H.A., AlJubab, H.A., Al-Hussaini, A.A., AlManjomi, F., Alsaad, A.B., Alsaleem, B., Andijani, A.A., Asery, A., Ballourah, W., Bleeker-Rovers, C.P., van Deuren, M., van der Flier, M., Gerkes, E.H., Gilissen, C., Habazi, M.K., Hehir-Kwa, J.Y., Henriët, S.S., Hoppenreijds, E.P., Hortillosa, S., Kerkhofs, C.H., Keski-Filppula, R., Lelieveld, S.H., Lone, K., MacKenzie, M.A., Mensenkamp, A.R., Moilanen, J., Nelen, M., Ten Oever, J., Potjewijd, J., van Paassen, P., Schuurs-Hoeijmakers, J.H.M., Simon, A., Stokowy, T., van de Vorst, M., Vreeburg, M., Wagner, A., van Well, G.T.J., Zafeiropoulou, D., Zonneveld-Huijssoon, E., Veltman, J.A., van Zelst-Stams, W.A.G., Fageih, E.A., van de Veerdonk, F.L., Netea, M.G., Hoischen, A., 2019. Exome sequencing in routine diagnostics: a generic test for 254 patients with primary immunodeficiencies. *Genome Med.* 11, 38. <https://doi.org/10.1186/s13073-019-0649-3>
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. <https://doi.org/10.1038/75556>
- Audano, P.A., Sulovari, A., Graves-Lindsay, T.A., Cantsilieris, S., Sorensen, M., Welch, A.E., Dougherty, M.L., Nelson, B.J., Shah, A., Dutcher, S.K., Warren, W.C., Magrini, V., McGrath, S.D., Li, Y.I., Wilson, R.K., Eichler, E.E., 2019. Characterizing the Major Structural Variant Alleles of the Human Genome. *Cell*

- Becker, J., Schwaab, R., Möller-Taube, A., Schwaab, U., Schmidt, W., Brackmann, H.H., Grimm, T., Olek, K., Oldenburg, J., 1996. Characterization of the factor VIII defect in 147 patients with sporadic hemophilia A: family studies indicate a mutation type-dependent sex ratio of mutation frequencies. *Am. J. Hum. Genet.* 58, 657–70.
- Cho, J.H., Abraham, C., 2007. Inflammatory Bowel Disease Genetics: Nod2. *Annu. Rev. Med.* 58, 401–416. <https://doi.org/10.1146/annurev.med.58.061705.145024>
- Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., Chan, A.P., 2012. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* 7, e46688. <https://doi.org/10.1371/journal.pone.0046688>
- Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., Coban Akdemir, Z.H., Doheny, K., Scott, A.F., Avramopoulos, D., Chakravarti, A., Hoover-Fong, J., Mathews, D., Witmer, P.D., Ling, H., Hetrick, K., Watkins, L., Patterson, K.E., Reinier, F., Blue, E., Muzny, D., Kircher, M., Bilguvar, K., López-Giráldez, F., Sutton, V.R., Tabor, H.K., Leal, S.M., Gunel, M., Mane, S., Gibbs, R.A., Boerwinkle, E., Hamosh, A., Shendure, J., Lupski, J.R., Lifton, R.P., Valle, D., Nickerson, D.A., Genomics, C. for M., Bamshad, M.J., 2015. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am. J. Hum. Genet.* 97, 199. <https://doi.org/10.1016/J.AJHG.2015.06.009>
- Ciccarelli, F., De Martinis, M., Ginaldi, L., 2014. An update on autoinflammatory diseases. *Curr. Med. Chem.* 21, 261–9. <https://doi.org/10.2174/09298673113206660303>
- Cooper, G.M., Stone, E.A., Asimenos, G., Program, S., Green, E.D., Batzoglou, S., Sidow, A., 2005. Distribution and intensity of constraint in mammalian genomic sequence.
- Courtois, G., Smahi, A., Reichenbach, J., Döffinger, R., Cancrini, C., Bonnet, M., Puel, A., Chable-Bessia, C., Yamaoka, S., Feinberg, J., Dupuis-Girod, S., Bodemer, C., Livadiotti, S., Novelli, F., Rossi, P., Fischer, A., Israël, A., Munnich, A., Le Deist, F., Casanova, J.-L., 2003. A hypermorphic IkappaBalpha mutation is associated with autosomal dominant anhidrotic ectodermal dysplasia and T cell immunodeficiency. *J. Clin. Invest.* 112, 1108–15. <https://doi.org/10.1172/JCI18714>
- de Valles-Ibáñez, G., Esteve-Solé, A., Piquer, M., González-Navarro, E.A., Hernandez-Rodriguez, J., Laayouni, H., González-Roca, E., Plaza-Martin, A.M., Deyà-Martínez, Á., Martín-Nalda, A., Martínez-Gallo, M., García-Prat, M., Del Pino-Molina, L., Cuscó, I., Codina-Solà, M., Batlle-Masó, L., Solís-Moruno, M., Marquès-Bonet, T., Bosch, E., López-Granados, E., Aróstegui, J.I., Soler-Palacín, P., Colobran, R., Yagüe, J., Alsina, L., Juan, M., Casals, F., 2018. Evaluating the Genetics of Common Variable Immunodeficiency: Monogenetic Model and Beyond. *Front. Immunol.* 9, 636. <https://doi.org/10.3389/fimmu.2018.00636>
- DiDonato, J., Mercurio, F., Rosette, C., Wu-Li, J., Suyang, H., Ghosh, S., Karin, M., 1996. Mapping of the inducible IkappaB phosphorylation sites that signal its ubiquitination and degradation. *Mol. Cell. Biol.* 16, 1295–1304. <https://doi.org/10.1128/MCB.16.4.1295>

- Fromer, M., Moran, J.L., Chambert, K., Banks, E., Bergen, S.E., Ruderfer, D.M., Handsaker, R.E., McCarroll, S.A., O'Donovan, M.C., Owen, M.J., Kirov, G., Sullivan, P.F., Hultman, C.M., Sklar, P., Purcell, S.M., 2012. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am. J. Hum. Genet.* 91, 597–607.
<https://doi.org/10.1016/j.ajhg.2012.08.005>
- Fromer, M., Purcell, S.M., 2014. Using XHMM Software to Detect Copy Number Variation in Whole-Exome Sequencing Data, in: *Current Protocols in Human Genetics*. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 7.23.1-7.23.21.
<https://doi.org/10.1002/0471142905.hg0723s81>
- Fuentes Fajardo, K. V., Adams, D., Mason, C.E., Sincan, M., Tifft, C., Toro, C., Boerkoel, C.F., Gahl, W., Markello, T., Markello, T., 2012. Detecting false-positive signals in exome sequencing. *Hum. Mutat.* 33, 609–613.
<https://doi.org/10.1002/humu.22033>
- Gallego-Bustos, F., Gotea, V., Ramos-Amador, J.T., Rodríguez-Pena, R., Gil-Herrera, J., Sastre, A., Delmiro, A., Rai, G., Elnitski, L., González-Granado, L.I., Allende, L.M., 2016. A case of IL-7R deficiency caused by a novel synonymous mutation and implications for mutation screening in SCID diagnosis. *Front. Immunol.* 7.
<https://doi.org/10.3389/fimmu.2016.00443>
- Gelfman, S., Wang, Q., McSweeney, K.M., Ren, Z., La Carpia, F., Halvorsen, M., Schoch, K., Ratzon, F., Heinzen, E.L., Boland, M.J., Petrovski, S., Goldstein, D.B., 2017. Annotating pathogenic non-coding variants in genic regions. *Nat. Commun.* 8, 236. <https://doi.org/10.1038/s41467-017-00141-2>
- Hayden, M.S., Ghosh, S., 2008. Shared Principles in NF- κ B Signaling. *Cell* 132, 344–362. <https://doi.org/10.1016/j.cell.2008.01.020>
- Hoffman, H.M., Broderick, L., 2016. Clinical reviews in allergy and immunology The role of the inflammasome in patients with autoinflammatory diseases.
<https://doi.org/10.1016/j.jaci.2016.05.001>
- Hoyos-Bachiloglu, R., Chou, J., Sodroski, C.N., Beano, A., Bainter, W., Angelova, M., Al Idrissi, E., Habazi, M.K., Alghamdi, H.A., Almanjomi, F., Al Shehri, M., Elsidig, N., Eldin, M.A., Knipe, D.M., AlZahrani, M., Geha, R.S., 2017. A digenic human immunodeficiency characterized by IFNAR1 and IFNGR2 mutations. *J. Clin. Invest.* 127, 4415–4420. <https://doi.org/10.1172/JCI93486>
- Huang, N., Lee, I., Marcotte, E.M., Hurles, M.E., 2010. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* 6, 1–11.
<https://doi.org/10.1371/journal.pgen.1001154>
- Hugot, J.P., Chamaillard, M., Zouali, H., Lesage, S., Cézard, J.P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C.A., Gassull, M., Binder, V., Finkel, Y., Cortot, A., Modigliani, R., Laurent-Puig, P., Gower-Rousseau, C., Macry, J., Colombel, J.F., Sahbatou, M., Thomas, G., 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411, 599–603.
<https://doi.org/10.1038/35079107>
- Karacan, İ., Balamir, A., Uğurlu, S., Aydın, A.K., Everest, E., Zor, S., Önen, M.Ö., Daşdemir, S., Özkaya, O., Sözeri, B., Tufan, A., Yıldırım, D.G., Yüksel, S., Ayaz, N.A., Ömeroğlu, R.E., Öztürk, K., Çakan, M., Söylemezoğlu, O., Şahin, S., Barut,

- K., Adroviç, A., Seyahi, E., Özdoğan, H., Kasapçopur, Ö., Turanlı, E.T., 2019. Diagnostic utility of a targeted next-generation sequencing gene panel in the clinical suspicion of systemic autoinflammatory diseases: a multi-center study. *Rheumatol. Int.* 39, 911–919. <https://doi.org/10.1007/s00296-019-04252-5>
- Khurana, E., Fu, Y., Chen, J., Gerstein, M., 2013. Interpretation of Genomic Variants Using a Unified Biological Network Approach. *PLoS Comput. Biol.* 9, e1002886. <https://doi.org/10.1371/journal.pcbi.1002886>
- Kircher, M., Witten, D.M., Jain, P., O’Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46, 310–5. <https://doi.org/10.1038/ng.2892>
- Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. <https://doi.org/10.1101/gr.129684.111>
- Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081. <https://doi.org/10.1038/nprot.2009.86>
- Lek, M., Karczewski, K.J., Minikel, E. V., Samocha, K.E., Banks, E., Fennell, T., O’Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. <https://doi.org/10.1038/nature19057>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B., Albers, C.A., Zhang, Z.D., Conrad, D.F., Lunter, G., Zheng, H., Ayub, Q., DePristo, M.A., Banks, E., Hu, M., Handsaker, R.E., Rosenfeld, J.A., Fromer, M., Jin, M., Mu, X.J., Khurana, E., Ye, K., Kay, M., Saunders, G.I., Suner, M.-M., Hunt, T., Barnes, I.H.A., Amid, C., Carvalho-Silva, D.R., Bignell, A.H., Snow, C., Yngvadottir, B., Bumpstead, S., Cooper, D.N., Xue, Y., Romero, I.G., Wang, J., Li, Y., Gibbs, R.A., McCarroll, S.A., Dermitzakis, E.T., Pritchard, J.K., Barrett, J.C., Harrow, J., Hurles, M.E., Gerstein, M.B., Tyler-Smith, C., Tyler-Smith, Chris, 2012. A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes. *Science* (80-.). 335, 823–828. <https://doi.org/10.1126/science.1215040>

- MacArthur, D.G., Tyler-Smith, C., 2010. Loss-of-function variants in the genomes of healthy humans. *Hum. Mol. Genet.* 19, R125-30.
<https://doi.org/10.1093/hmg/ddq365>
- Martínez Valle, F., Balada, E., Ordi-Ros, J., Vilardell-Tarres, M., 2008. DNase 1 and systemic lupus erythematosus. *Autoimmun. Rev.* 7, 359–363.
<https://doi.org/10.1016/j.autrev.2008.02.002>
- Martorana, D., Bonatti, F., Mozzoni, P., Vaglio, A., Percesepe, A., 2017. Monogenic Autoinflammatory Diseases with Mendelian Inheritance: Genes, Mutations, and Genotype/Phenotype Correlations. *Front. Immunol.* 8, 344.
<https://doi.org/10.3389/fimmu.2017.00344>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M., 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome R* 20, 1297–303.
- Mensa-Vilaró, A., Bravo García-Morato, M., de la Calle-Martin, O., Franco-Jarava, C., Martínez-Saavedra, M.T., González-Granado, L.I., González-Roca, E., Fuster, J.L., Alsina, L., Mutchinick, O.M., Balderrama-Rodríguez, A., Ramos, E., Modesto, C., Mesa-del-Castillo, P., Ortego-Centeno, N., Clemente, D., Souto, A., Palmou, N., Remesal, A., Leslie, K.S., Gómez de la Fuente, E., Yadira Bravo Gallego, L., Campistol, J.M., Dhouib, N.G., Bejaoui, M., Dutra, L.A., Terreri, M.T., Mosquera, C., González, T., Cañellas, J., García-Ruiz de Morales, J.M., Wouters, C.H., Bosque, M.T., Cham, W.T., Jiménez-Treviño, S., de Inocencio, J., Bloomfield, M., Pérez de Diego, R., Martínez-Pomar, N., Rodríguez-Pena, R., González-Santesteban, C., Soler-Palacín, P., Casals, F., Yagüe, J., Allende, L.M., Rodríguez-Gallego, J.C., Colobran, R., Martínez-Martínez, L., López-Granados, E., Aróstegui, J.I., 2019. Unexpected relevant role of gene mosaicism in patients with primary immunodeficiency diseases. *J. Allergy Clin. Immunol.* 143, 359–368.
<https://doi.org/10.1016/j.jaci.2018.09.009>
- Meyts, I., Bosch, B., Bolze, A., Boisson, B., Itan, Y., Belkadi, A., Pedergrana, V., Moens, L., Picard, C., Cobat, A., Bossuyt, X., Abel, L., Casanova, J.L., 2016. Exome and genome sequencing for inborn errors of immunity. *J. Allergy Clin. Immunol.* 138, 957–969. <https://doi.org/10.1016/j.jaci.2016.08.003>
- Negróni, A., Pierdomenico, M., Cucchiara, S., Stronati, L., 2018. NOD2 and inflammation: current insights. *J. Inflamm. Res.* Volume 11, 49–60.
<https://doi.org/10.2147/JIR.S137606>
- Notarangelo, L.D., Sorensen, R., 2008. Is it necessary to identify molecular defects in primary immunodeficiency disease? *J. Allergy Clin. Immunol.* 122, 1069–1073.
<https://doi.org/10.1016/j.jaci.2008.08.038>
- Omoyinmi, E., Standing, A., Keylock, A., Price-Kuehne, F., Melo Gomes, S., Rowczenio, D., Nanthapaisal, S., Cullup, T., Nyanhete, R., Ashton, E., Murphy, C., Clarke, M., Ahlfors, H., Jenkins, L., Gilmour, K., Eleftheriou, D., Lachmann, H.J., Hawkins, P.N., Klein, N., Brogan, P.A., 2017. Clinical impact of a targeted next-generation sequencing gene panel for autoinflammation and vasculitis. *PLoS One* 12, e0181874. <https://doi.org/10.1371/journal.pone.0181874>
- Papa, R., Rusmini, M., Volpi, S., Caorsi, R., Picco, P., Grossi, A., Caroli, F., Bovis, F., Musso, V., Obici, L., Castana, C., Ravelli, A., Van Gijn, M.E., Ceccherini, I.,

- Gattorno, M., 2019. Next generation sequencing panel in undifferentiated autoinflammatory diseases identifies patients with colchicine-responder recurrent fevers. *Rheumatology*. <https://doi.org/10.1093/rheumatology/kez270>
- Petrovski, S., Gussow, A.B., Wang, Q., Halvorsen, M., Han, Y., Weir, W.H., Allen, A.S., Goldstein, D.B., 2015. The Intolerance of Regulatory Sequence to Genetic Variation Predicts Gene Dosage Sensitivity. *PLOS Genet.* 11, e1005492. <https://doi.org/10.1371/journal.pgen.1005492>
- Picard, C., Bobby Gaspar, H., Al-Herz, W., Bousfiha, A., Casanova, J.-L., Chatila, T., Crow, Y.J., Cunningham-Rundles, C., Etzioni, A., Franco, J.L., Holland, S.M., Klein, C., Morio, T., Ochs, H.D., Oksenhendler, E., Puck, J., Tang, M.L.K., Tangye, S.G., Torgerson, T.R., Sullivan, K.E., 2018. International Union of Immunological Societies: 2017 Primary Immunodeficiency Diseases Committee Report on Inborn Errors of Immunity. *J. Clin. Immunol.* 38, 96–128. <https://doi.org/10.1007/s10875-017-0464-9>
- Posey, J.E., Rosenfeld, J.A., James, R.A., Bainbridge, M., Niu, Z., Wang, X., Dhar, S., Wiszniewski, W., Akdemir, Z.H.C., Gambin, T., Xia, F., Person, R.E., Walkiewicz, M., Shaw, C.A., Sutton, V.R., Beaudet, A.L., Muzny, D., Eng, C.M., Yang, Y., Gibbs, R.A., Lupski, J.R., Boerwinkle, E., Plon, S.E., 2016. Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genet. Med.* 18, 678–685. <https://doi.org/10.1038/gim.2015.142>
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L., ACMG Laboratory Quality Assurance Committee, 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 17, 405–24. <https://doi.org/10.1038/gim.2015.30>
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26. <https://doi.org/10.1038/nbt.1754>
- Rodero, M.P., Tesser, A., Bartok, E., Rice, G.I., Della Mina, E., Depp, M., Beitz, B., Bondet, V., Cagnard, N., Duffy, D., Dussiot, M., Frémond, M.-L., Gattorno, M., Guillem, F., Kitabayashi, N., Porcheray, F., Rieux-Laucat, F., Seabra, L., Uggenti, C., Volpi, S., Zeef, L.A.H., Alyanakian, M.-A., Beltrand, J., Bianco, A.M., Boddaert, N., Brouzes, C., Candon, S., Caorsi, R., Charbit, M., Fabre, M., Faletra, F., Girard, M., Harroche, A., Hartmann, E., Lasne, D., Marcuzzi, A., Neven, B., Nitschke, P., Pascreau, T., Pastore, S., Picard, C., Picco, P., Piscianz, E., Polak, M., Quartier, P., Rabant, M., Stocco, G., Taddio, A., Uettwiller, F., Valencic, E., Vozzi, D., Hartmann, G., Barchet, W., Hermine, O., Bader-Meunier, B., Tommasini, A., Crow, Y.J., 2017. Type I interferon-mediated autoinflammation due to DNase II deficiency. *Nat. Commun.* 8, 2176. <https://doi.org/10.1038/s41467-017-01932-3>
- RStudio Team, 2016. RStudio: Integrated development for R [WWW Document]. URL <https://www.rstudio.com/> (accessed 6.27.19).
- Seleman, M., Hoyos-Bachiloglu, R., Geha, R.S., Chou, J., 2017. Uses of next-generation sequencing technologies for the diagnosis of primary immunodeficiencies. *Front. Immunol.* <https://doi.org/10.3389/fimmu.2017.00847>

- Sieni, E., Cetica, V., Santoro, A., Beutel, K., Mastrodicasa, E., Meeths, M., Ciambotti, B., Brugnolo, F., zur Stadt, U., Pende, D., Moretta, L., Griffiths, G.M., Henter, J.-I., Janka, G., Arico, M., 2011. Genotype-phenotype study of familial haemophagocytic lymphohistiocytosis type 3. *J. Med. Genet.* 48, 343–352. <https://doi.org/10.1136/jmg.2010.085456>
- Simons, C., Rash, L.D., Crawford, J., Ma, L., Cristofori-Armstrong, B., Miller, D., Ru, K., Baillie, G.J., Alanay, Y., Jacquinet, A., Debray, F.-G., Verloes, A., Shen, J., Yesil, G., Guler, S., Yuksel, A., Cleary, J.G., Grimmond, S.M., McGaughran, J., King, G.F., Gabbett, M.T., Taft, R.J., 2015. Mutations in the voltage-gated potassium channel gene *KCNH1* cause Temple-Baraitser syndrome and epilepsy. *Nat. Genet.* 47, 73–77. <https://doi.org/10.1038/ng.3153>
- Steinberg, J., Honti, F., Meader, S., Webber, C., 2015. Haploinsufficiency predictions without study bias. *Nucleic Acids Res.* 43, e101–e101. <https://doi.org/10.1093/nar/gkv474>
- Stray-Pedersen, A., Sorte, H.S., Samarakoon, P., Gambin, T., Chinn, I.K., Coban Akdemir, Z.H., Erichsen, H.C., Forbes, L.R., Gu, S., Yuan, B., Jhangiani, S.N., Muzny, D.M., Rødningen, O.K., Sheng, Y., Nicholas, S.K., Noroski, L.M., Seeborg, F.O., Davis, C.M., Canter, D.L., Mace, E.M., Vece, T.J., Allen, C.E., Abhyankar, H.A., Boone, P.M., Beck, C.R., Wiszniewski, W., Fevang, B., Aukrust, P., Tjønnfjord, G.E., Gedde-Dahl, T., Hjorth-Hansen, H., Dybedal, I., Nordøy, I., Jørgensen, S.F., Abrahamsen, T.G., Øverland, T., Bechensteen, A.G., Skogen, V., Osnes, L.T.N., Kulseth, M.A., Prescott, T.E., Rustad, C.F., Heimdal, K.R., Belmont, J.W., Rider, N.L., Chinen, J., Cao, T.N., Smith, E.A., Caldirola, M.S., Bezrodnik, L., Lugo Reyes, S.O., Espinosa Rosales, F.J., Guerrero-Cursaru, N.D., Pedroza, L.A., Poli, C.M., Franco, J.L., Trujillo Vargas, C.M., Aldave Becerra, J.C., Wright, N., Issekutz, T.B., Issekutz, A.C., Abbott, J., Caldwell, J.W., Bayer, D.K., Chan, A.Y., Aiuti, A., Cancrini, C., Holmberg, E., West, C., Burstedt, M., Karaca, E., Yesil, G., Artac, H., Bayram, Y., Atik, M.M., Eldomery, M.K., Ehlayel, M.S., Jolles, S., Flatø, B., Bertuch, A.A., Hanson, I.C., Zhang, V.W., Wong, L.J., Hu, J., Walkiewicz, M., Yang, Y., Eng, C.M., Boerwinkle, E., Gibbs, R.A., Shearer, W.T., Lyle, R., Orange, J.S., Lupski, J.R., 2017. Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *J. Allergy Clin. Immunol.* 139, 232–245. <https://doi.org/10.1016/j.jaci.2016.05.042>
- Suzuki, T., Sasahara, Y., Kikuchi, A., Kakuta, H., Kashiwabara, T., Ishige, T., Nakayama, Y., Tanaka, M., Hoshino, A., Kanegane, H., Abukawa, D., Kure, S., 2017. Targeted Sequencing and Immunological Analysis Reveal the Involvement of Primary Immunodeficiency Genes in Pediatric IBD: a Japanese Multicenter Study. *J. Clin. Immunol.* 37, 67–79. <https://doi.org/10.1007/s10875-016-0339-5>
- The Gene Ontology Consortium, 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330–D338. <https://doi.org/10.1093/nar/gky1055>
- Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38, e164–e164. <https://doi.org/10.1093/nar/gkq603>
- Wekell, P., Berg, S., Karlsson, A., Fasth, A., 2017. Toward an inclusive, congruent, and precise definition of autoinflammatory diseases. *Front. Immunol.* 8, 6–10. <https://doi.org/10.3389/fimmu.2017.00497>

- Wekell, P., Karlsson, A., Berg, S., Fasth, A., Wekell, C.P., 2016. Review of autoinflammatory diseases, with a special focus on periodic fever, aphthous stomatitis, pharyngitis and cervical adenitis syndrome. <https://doi.org/10.1111/apa.13531>
- Yao, Q., Lacbawan, F., Li, J., 2016. Adult autoinflammatory disease frequency and our diagnostic experience in an adult autoinflammatory clinic. *Semin. Arthritis Rheum.* 45, 633–637. <https://doi.org/10.1016/j.semarthrit.2015.10.012>
- Yasutomo, K., Horiuchi, T., Kagami, S., Tsukamoto, H., Hashimura, C., Urushihara, M., Kuroda, Y., 2001. Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat. Genet.* 28, 313–314. <https://doi.org/10.1038/91070>
- Yska, H.A.F., Elsink, K., Kuijpers, T.W., Frederix, G.W.J., van Gijn, M.E., van Montfrans, J.M., 2019. Diagnostic Yield of Next Generation Sequencing in Genetically Undiagnosed Patients with Primary Immunodeficiencies: a Systematic Review. *J. Clin. Immunol.* 39, 577–591. <https://doi.org/10.1007/s10875-019-00656-x>