# Activating germline mutations in STAT3 cause early-onset multi-organ autoimmune disease.

Flanagan, Sarah E; Haapaniemi, Emma; Russell, Mark A; Caswell, Richard; Lango Allen, Hana; De Franco, Elisa; McDonald, Timothy J; Rajala, Hanna; Ramelius, Anita; Barton, John; Heiskanen, Kaarina; Heiskanen-Kosma, Tarja; Kajosaari, Merja; Murphy, Nuala P; Milenkovic, Tatjana; Seppänen, Mikko; Lernmark, Åke; Mustjoki, Satu; Otonkoski, Timo; Kere, Juha; Morgan, Noel G; Ellard, Sian; Hattersley, Andrew T

*Published in:*
Nature Genetics

*DOI:*
10.1038/ng.3040

2014

# Activating germline mutations in *STAT3* cause early-onset multi-organ autoimmune disease

Sarah E. Flanagan,[1,16] Emma Haapaniemi,[2,3,16] Mark A. Russell,[1,16] Richard Caswell,[1] Hana Lango Allen,[1] Elisa De Franco,[1] Timothy J. McDonald,[1] Hanna Rajala,[4,5] Anita Ramelius,[6,7] John Barton,[8] Kaarina Heiskanen,[3,9] Tarja Heiskanen-Kosma,[10] Merja Kajosaari,[9] Nuala P. Murphy,[11] Tatjana Milenkovic,[12] Mikko Seppänen,[13] Åke Lernmark,[6,7] Satu Mustjoki,[4,5] Timo Otonkoski,[3,9] Juha Kere,[2,3,14,15] Noel G. Morgan,[1] Sian Ellard[1] and Andrew T. Hattersley[1]

[1] Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, EX2 5DW, UK
[2] Folkhälsan Institute of Genetics, University of Helsinki, Helsinki, Finland
[3] Research Programs Unit, Molecular Neurology, University of Helsinki, Helsinki, Finland
[4] Hematology Research Unit Helsinki, Department of Hematology, University of Helsinki
[5] Helsinki University Central Hospital Cancer Center, Helsinki, Finland
[6] Department of Clinical Sciences, Lund University, Lund, Sweden
[7] CRC, Skåne University Hospital SUS, Malmö, Sweden
[8] Bristol Royal Hospital for Children, Upper Maudlin Street, Bristol, BS2 8BJ, UK
[9] Children's Hospital, Helsinki University Central Hospital, Helsinki, Finland
[10] Department of Pediatrics, Kuopio University Hospital, Helsinki, Finland
[11] Department of Diabetes and Endocrinology, Children's University Hospital, Temple St., Dublin 1, Ireland
[12] Department of Endocrinology, Institute for Mother and Child Health Care of Serbia 'Dr Vukan Cupic', Belgrade, Serbia
[13] Immunodeficiency Unit, Division of Infectious Diseases, Helsinki University Central Hospital, Helsinki, Finland
[14] Department of Biosciences and Nutrition, Karolinska Institutet, Hälsovägen 7, 14183 Huddinge, Sweden
[15] Center for Innovative Medicine, Karolinska Institutet, Hälsovägen 7, 14183 Huddinge, Sweden
[16] These authors contributed equally to this work

**Corresponding Authors:**

**Prof Andrew T. Hattersley**
University of Exeter Medical School
RILD Building,
Barrack Road
Exeter, EX2 5DW, UK
A.Hattersley@exeter.ac.uk
+44 (0)1392 408260

**Prof Noel G. Morgan**
University of Exeter Medical School
RILD Building,
Barrack Road
Exeter, EX2 5DW, UK
N.G.Morgan@exeter.ac.uk
+44 (0)1392 408300

**Prof Juha Kere**
Folkhälsan Institute of Genetics, and Research Programs Unit, Molecular Neurology, University of Helsinki, Helsinki, Finland
juha.kere@ki.se
+46734213550

**Monogenic causes of autoimmunity give key insights to the complex regulation of the immune system. We report a new monogenic cause of autoimmunity resulting from *de novo* germline activating *STAT3* mutations in 5 individuals with a spectrum of early-onset autoimmune disease including type 1 diabetes. These findings emphasise the critical role of STAT3 in autoimmune disease and contrast with the germline inactivating *STAT3* mutations that result in Hyper IgE syndrome.**

Autoimmune disorders are usually multifactorial in aetiology, involving a combination of the background risk due to an individual's genetic make-up and environmental exposure. Considerable advances in understanding the genetic susceptibility to autoimmune disease have identified common and disease specific polymorphisms in the Human Leukocyte Antigen (HLA) region and throughout the genome.[1-3] Rarely monogenic defects can explain specific clustering of autoimmune conditions and give important biological insights.[4-6]

Type 1 diabetes (T1D) can occur in association with other autoimmune conditions because of a shared polygenic predisposition or rarely as part of a monogenic polyautoimmune disorder. T1D is a common feature of Immunodysregulation, Polyendocrinopathy, Enteropathy X-linked (IPEX) syndrome due to *FOXP3* mutations and Autoimmune Polyendocrinopathy Syndrome 1 (APS1) due to recessive mutations in *AIRE*.[4,5] In these polyendocrinopathy syndromes the autoimmune disease has a very young age of onset with T1D usually diagnosed before 3 months in IPEX syndrome which is in contrast to polygenic T1D that is very rare before 6 months. [7] Investigation of individuals with early-onset multiple autoimmune features may therefore reveal novel monogenic causes of autoimmunity.

We initially searched for a *de novo* mutation in an individual with early-onset polyautoimmunity (T1D (2wks), autoimmune hypothyroidism (3yrs), coeliac disease (17months)), by exome sequencing and comparison with variants identified in the unaffected parents. Heterozygous variants called using the Genome Analysis Tool Kit were filtered by removing non-coding or synonymous variants, those in dbSNP131 or 1000 Genomes Project databases and variants identified in either parent (**Online Methods**). This reduced the number of potentially pathogenic *de novo* mutations to a single novel heterozygous missense substitution, p.Thr716Met (c.2147C>T), in the transactivation domain of STAT3 (**Supplementary Table 1**).

Sequencing of the *STAT3* coding exons was performed in i) 24 individuals with >2 early-onset (diagnosis <5 years) autoimmune disorders of unknown cause and ii) 39 individuals with

isolated permanent diabetes diagnosed <6 months, aged <5 years at referral, and with 20 known causes of neonatal diabetes excluded (**Online Methods, Supplementary Table 2**). We identified 3 different heterozygous missense substitutions (p.K392R, p.N646K, p.K658N) in 4 individuals (3 polyautoimmune disease and 1 isolated permanent neonatal diabetes aged 3 yrs). Analysis of parental DNA samples established that all mutations had arisen *de novo* (**Online Methods**). With the exception of p.K658N, which has been reported as a somatic mutation causing Large Granular Lymphocytic (LGL) Leukaemia[8], all mutations were novel and absent in 1000 Genomes and NHLBI exome databases. All substitutions affect highly conserved residues within the SH2, transactivation or DNA binding domains (**Fig 1a**). The finding of a spontaneous mutation in 5 individuals with early-onset autoimmunity provides overwhelming genetic evidence that these are disease causing mutations.

The activity of each mutation was assessed using a STAT3-responsive dual luciferase reporter assay (**Online Methods**). The 4 different *STAT3* mutations identified in individuals with polyautoimmune disease, 2 dominant negative Hyper IgE mutations, and wild-type protein were generated and transiently transfected into cultured HEK293 cells. Under non-stimulated (basal) conditions all 4 novel *STAT3* mutations resulted in an increase in reporter activity compared to wild-type or the Hyper IgE inactivating mutations (**Fig 1b**). Three of the 4 mutations also showed a less marked increase relative to wild-type when IL-6 stimulated: for the p.N646K mutation, the least activating of the mutations tested, although there is a significant increase in reporter activity under non-stimulated (basal) conditions there is not an increase above wild type when IL-6 stimulated (**Fig 1b**). To establish underlying cytokine-related gain-of-function we performed T-cell immunophenotyping and activation assays for two patients (mutations p.K658N and p.K392R). Both showed reduced regulatory T-cell numbers, and CD4+ cells of patient 2 also showed increased cytokine production (IFN-γ and TNF) when stimulated (**Supplementary Table 3 and Supplementary Figure 3**). This provides evidence for enhanced Th1 response *in vivo*. These functional studies support that all of the mutations causing polyautoimmune disease are activating and lead to increased basal *STAT3* activity *in vitro*.

The diabetes (4/5 individuals) presented early (2.5 [0-43] weeks (median [range])) and was insulin treated from diagnosis in doses required for total endogenous insulin deficiency. Three of these 4 patients had detectable islet autoantibodies for T1D (**Supplementary Table 3**) supporting an autoimmune aetiology. Autoimmune destruction of beta cells starting during fetal life is suggested by intra-uterine growth retardation, a likely consequence of reduced

insulin secretion *in utero*, and early-onset diabetes (≤3 weeks). Additional autoimmune conditions diagnosed during childhood included autoimmune enteropathy, autoimmune interstitial lung disease, juvenile-onset arthritis and primary hypothyroidism. Other common features were short stature (5/5 (<2SDS)) and eczema (4/5) (**Supplementary Table 3**). The young age at diagnosis of the autoimmune features is consistent with *STAT3* mutations causing accelerated autoimmune disease.

The Signal Transducer and Activator of Transcription (STAT) factors are involved in multiple processes, including early development, cellular proliferation, survival and differentiation. It is possible that STAT3 activation leads to autoimmunity through impairing the development of regulatory T-cells and promoting the expansion and activation of Th17 cells.[9,10] Th17 expansion, activation and dysregulation are thought to play a critical role in many autoimmune diseases including Type 1 diabetes.[11,12] In addition, direct cell-type specific effects of constitutively active STAT3 may also play a role in autoimmunity development.

Germline dominant-negative *STAT3* mutations in patients with Hyper-IgE syndrome cause a primary immunodeficiency disorder characterised by elevated serum IgE and recurrent staphylococcal infections due to deficiency of Th17 cells.[12,13] The multiple autoimmune features observed in our cohort with germline activating *STAT3* mutations is in contrast to the severe immunodeficiency observed in individuals with the Hyper IgE syndrome (**Fig 2a**).

Somatic activating *STAT3* mutations have been reported in 40% of individuals with LGL leukaemia[8] and 8% of inflammatory hepatocellular adenomas.[14] The primary presentation of LGL-leukaemia is recurrent bacterial infections, fatigue, and autoimmune cytopenias typically in the 6$^{th}$ decade of life. These individuals frequently (26%) have adult-onset rheumatoid arthritis but diabetes or other autoimmune disease is not reported.[8] Therefore different phenotypes result from somatic and germline mutations

Common variation in *STAT3* is associated with autoimmune disease with single nucleotide polymorphisms conferring altered susceptibility to Crohn's disease[2], psoriasis[15]; and multiple sclerosis[16] but interestingly these conditions were not seen in our patient series. Suggestive associations were seen in T1D and rheumatoid arthritis but did not reach genome-wide significance.[17,18] Further studies investigating the role of rare-coding variants in susceptibility to common organ-specific autoimmune disease are warranted.

*STAT3* associated polyautoimmunity has a different, but overlapping, clinical phenotype from the 4 previously described monogenic autoimmune syndromes (**Supplementary Table 3**). Whilst these disorders are extremely rare they offer valuable insights into the biology of the immune system in T1D and other related autoimmune disease. Further large scale sequencing efforts of individuals with multiple early-onset autoimmune disorders are required to identify novel monogenic conditions which will shed further light on the complex regulation of the adaptive immune response.

**Acknowledgements**

The authors thank John Chilton, Annet Damhuis, Rahul Raman and Bingmei Yang for technical assistance. This work was supported by NIHR Exeter Clinical Research Facility through funding for SE and ATH. SE and ATH are supported by Wellcome Trust Senior Investigator awards. Funding was provided by Diabetes UK and The Finnish Medical Foundation.

**Author Contributions**

S.E.F., S.E., and A.T.H. designed the study. N.P.H., T.M., T.O., E.H., K.H., T.H-K., M.K. and J.B. recruited patients to the study. R.C. and E.D.F. performed the exome sequencing and targeted next generation sequence analysis. H.L.A. did the bioinformatic analysis. S.E.F. and E.H. did the Sanger sequencing analysis and the interpretation of the resulting data. S.E.F., T.J.M., E.H. M.S., J.K. and A.T.H. analysed the clinical data. M.A.R. and N.G.M. designed and performed the functional studies. H.R. and S.M. performed T-cell assays. S.E.F., M.A.R. and A.T.H. prepared the draft manuscript. All authors contributed to discussion of the results and to manuscript preparation.

**Competing Financial Interests**

The authors declare no competing financial interests.

**Figure Legends**

1. **a)** Schematic representation of STAT3 adapted from[19]. The positions of the 4 different *de novo* mutations identified in 5 individuals with the polyautoimmune syndrome are shown below the STAT3 domains. The highly conserved sequence at the position of each mutation is provided for various species.
   **b)** STAT3 activity of polyautoimmune mutants under (upper) non-stimulated and (lower) IL-6 (20ng/ml) stimulated conditions. The STAT3 reporter activity of 4 polyautoimmune mutants (p.K392R, p.N646K, p.K658N, p.T716M) was examined alongside that of 2 previously described Hyper-IgE mutations (p.R382W, p.V637M)[19] and the wildtype (WT) following transient transfection into HEK293 cells. The dotted line indicates the activity of the WT under either basal (upper) or IL-6 stimulated (lower) conditions. Data are presented as an average fold change relative to WT (n=3) under each experimental condition ±SEM. Typically, IL-6 caused a 20-30-fold increase in activity above basal in cells transfected with WT STAT3.*p<0.05,**p<0.01, ***p<0.001.

**2.** Clinical characteristics associated with activating germline *STAT3* mutations causing the polyautoimmune syndrome and inactivating germline mutations causing Hyper IgE syndrome.[20]

## Online Methods

### Cohort selection and sample preparation

Twenty-five individuals with early-onset polyautoimmune disease (diagnosed before 5 years of age) and 39 subjects with isolated permanent diabetes diagnosed before 6 months were recruited by their clinicians for molecular genetic analysis in the Exeter Molecular Genetics Laboratory (n=63) or the Folkhälsan Institute of Genetics, University of Helsinki (n=1). Genomic DNA was extracted from peripheral leukocytes using standard procedures. All subjects and/or their parents gave informed consent for genetic testing and institutional review board approval was received for this study.

### Exome sequencing and variant calling

Genomic regions corresponding to NCBI Consensus Coding Sequence (CCDS) database were captured and amplified using Agilent's SureSelect Human All Exon Kit (v1). Paired-end sequencing was performed on an Illumina GAII, one lane per sample, 101 or 76bp read length. The resulting reads were aligned to the hg19 reference genome with BWA providing mean target coverage of 66.3 reads per base. At least 72% of the targeted bases were covered by at least 20 reads. Variants were called with GATK UnifiedGenotyper and annotated using Annovar and SeattleSeq Annotation server, as previously described.[21] Variant filtering steps are shown in **Supplementary Table 1**.

### *STAT3* sequencing and microsatellite analysis

Sanger sequencing was undertaken in patient 1 and her unaffected parents to confirm that the p.T716M *STAT3* variant had arisen *de novo*. Exons 2-24 and intron/exon boundaries of *STAT3* (NM_139276.2) were Sanger sequenced in a further 24 individuals with at least 2 early-onset autoimmune features of unknown cause (**Supplementary Table 2**). Primers for *STAT3* exons 2-24 are provided in **Supplementary Table 5**.

Targeted next-generation sequencing of *STAT3* was undertaken on a further 39 individuals with isolated permanent diabetes diagnosed before the age of 6 months of unknown cause without additional autoimmune features. All patients were less than 5 years of age at the time of genetic testing. We adapted our custom Agilent SureSelect exon-capture assay (Agilent Technologies, Santa Clara, CA, USA) to include baits for exons 2-24 and intron/exon boundaries of *STAT3* (sequences available on request to authors).[22] Samples were fragmented using a Bioruptor (Diagenode, Liège, Belgium), indexed for multiplexing and hybridised (in pools of 12 samples) according to the manufacturer's instructions. Sequencing was performed with an Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) (48 samples per lane) and 100 bp paired end reads. Data were processed to identify potential pathogenic mutations located within 50 bp upstream and 10 bp downstream of each exon.

We identified *STAT3* mutations in a further 3 individuals with early onset autoimmune disease (4 of 25, 16% of cohort) and 1 individual with permanent neonatal diabetes (1 of 39, 2.6% of cohort). This brought the total number of *STAT3* positive subjects to 5. Sanger sequencing of parental samples confirmed all mutations had arisen *de novo*. Biological relationships were confirmed by microsatellite analysis using the PowerPlex kit (PowerPlex 16 System, Promega, Southampton, UK).

In total four different mutations were identified in 5 unrelated individuals. All mutations affected residues in the highly conserved DNA binding domain (x1), SH2 domain (x2) and transactivation domain (x1) (conserved to Zebrafish). None of the mutations were present in dbSNP132, 1000 Genomes Project database (based on 1094 individuals) or the Exome Sequencing Project (based on 6500 individuals).

**Functional studies of *STAT3* mutations**

Mutations within human STAT3 (Source Bioscience) were generated using the QuickChange site-directed mutagenesis kit following the manufacturer's guidelines (Agilent Technology). The primer pairs used to generate each mutant are provided in **Supplementary table 6.** The success of all mutatagenesis reactions was confirmed by direct sequencing of the entire STAT3 insert (Eurofins). Following mutagenesis, STAT3 inserts were subcloned into the multiple cloning site of a pcDNA5/FRT/TO expression vector between AflII and EcoRV restriction sites.

The transcriptional activity of STAT3 was assessed via a STAT3 responsive dual firefly/Renilla luciferase Cignal reporter system (Qiagen). HEK293 cells were seeded at a density of 1 x $10^5$ cells/well, and were transfected after 24h with a combination of 200ng Cignal reporter assay constructs and 400ng WT or mutant STAT3 pcDNA5/FRT/TO using the Attractene transfection reagent according to the manufacturer's instructions (Qiagen). Cells were incubated in the transfection mix for 24h and, where appropriate, 20ng/ml IL-6 was also included for the final 18h. STAT3 reporter activity was assessed using a dual luciferase reporter assay system (Promega). To confirm that equivalent amounts of STAT3 protein were expressed following transfection of each construct, cells were lysed and protein extracted prior to Western blotting with anti-STAT3 antibody (Cell Signalling) as described previously.[23]

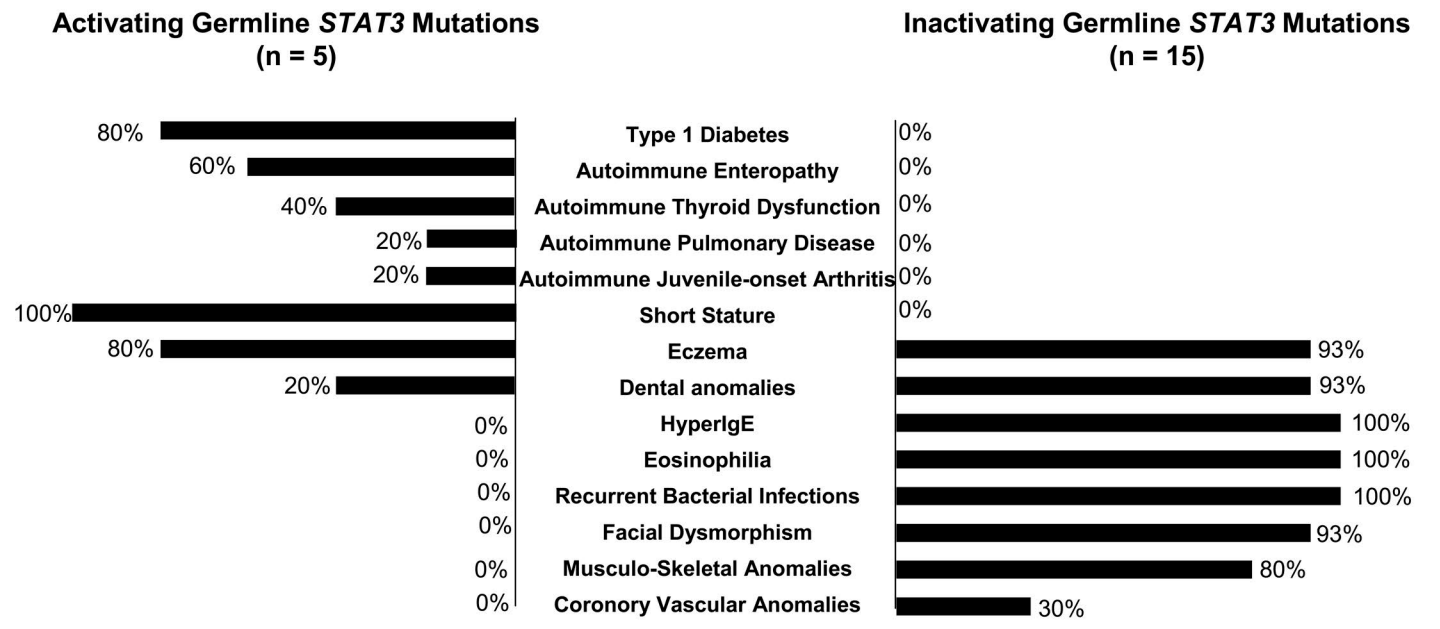
**Regulatory T-cell immunophenotyping and evaluation of T-cell cytokine production**

Fresh peripheral blood mononuclear cells (PBMNCs) from patients 2 and 5 (**Supplementary Table 3**) and six healthy controls were used. For regulatory T-cell immunophenotyping, cells were stained using monoclonal antibodies (mAb) against the antigens CD3, CD4 and CD25 (BD Biosciences) and FOXP3 (clone259D, eBioscience) and data collected with 4- or 6-color flow cytometry. Regulatory T-cells were defined as $CD3^+CD4^+CD25^{high}FOXP3^+$. Evaluation of T-cell cytokine production is described in detail elsewhere[23]. Briefly, fresh mononuclear cells were stimulated for 6 hours with anti-CD3, anti-CD28 and anti-CD49d (BD Biosciences). The cells were analyzed using a 6-color flow cytometry panel with mAbs against the antigens CD3, CD4, CD8, IFN-γ and TNF-α (BD Biosciences). The data were analyzed with FACSAria II flow cytometer and FACSDiva software (BD Biosciences).

**References**


**1** **Barrett, J. C. *et al. Nat Genet* 41, 703-707 (2009).**
**2** **Barrett, J. C. *et al. Nat Genet* 40, 955-962 (2008).**
**3** **Stahl, E. A. *et al. Nat Genet* 42, 508-514 (2010).**
**4** **Bennett, C. L. *et al. Nat Genet* 27, 20-21 (2001).**
**5** **Finnish-German APECED Consortium. *Nat Genet* 17, 399-403 (1997).**
**6** **Sharfe, N., *et al. Proc Natl Acad Sci U S A* 94, 3168-3171 (1997).**

7 Rubio-Cabezas, O. *et al. Diabetes Care* 32, 111-116 (2009).
8 Koskela, H. L. *et al. N Engl J Med* 366, 1905-1913 (2012).
9 Yang, X. O. *et al. J Biol Chem* 282, 9358-9363 (2007).
10 Harris, T. J. *et al. J Immunol* 179, 4313-4317 (2007).
11 Shao, S. *et al. Cell Immunol* 280, 16-21 (2012).
12 Ma, C. S. *et al. J Exp Med* 205, 1551-1557 (2008).
13 Minegishi, Y. *et al. Nature* 448, 1058-1062 (2007).
14 Pilati, C. *et al. J Exp Med* 208, 1359-1366 (2011).
15 Tsoi, L. C. *et al. Nat Genet* 44, 1341-1348 (2012).
16 Jakkula, E. *et al. Am J Hum Genet* 86, 285-291 (2010).
17 Fung, E. Y. *et al. Genes Immun* 10, 188-191 (2009).
18 Seddighzadeh, M. *et al. J Rheumatol* 39, 1509-1516, (2012).
19 Holland, S. M. *et al. N Engl J Med* 357, 1608-1619 (2007).
20 Heimall, J. *et al. Clin Immunol* 139, 75-84 (2011).
21 Lango Allen, H. *et al. Nat Genet* 44, 20-22 (2012).
22 Ellard, S. *et al. Diabetologia* 56, 1958-1963 (2013).
23 Russell, M. A., *et al. Islets* 5, 95-105 (2013).


Activating Germline STAT3 Mutations
(n = 5)
Inactivating Germline STAT3 Mutations
(n = 15)
80%
60%
40%
20%
20%
100%
80%
20%
0%
0%
0%
0%
0%
0%
Type 1 Diabetes
Autoimmune Enteropathy
Autoimmune Thyroid Dysfunction
Autoimmune Pulmonary Disease
Autoimmune Juvenile-onset Arthritis
Short Stature
Eczema
Dental anomalies
HyperIgE
Eosinophilia
Recurrent Bacterial Infections
Facial Dysmorphism
Musculo-Skeletal Anomalies
Coronory Vascular Anomalies
0%
0%
0%
0%
0%
0%
93%
93%
100%
100%
100%
93%
80%
30%