Supplementary Note

Patient case histories and identification of mutations

A 19-year-old Caucasian girl (subject A.II.1), with no significant family history (**Fig. 1a – left family**) presented to the department of hematology with severe anemia, thrombocytopenia, leukopenia, massive splenomegaly (**Fig. 1c – top left**), persistent fever greater than 40°C and myalgia. No infectious cause for her fever was identified, despite multiple blood and bone marrow cultures and bone marrow and gut biopsies. She did not respond to broad-spectrum antibiotics. A bone marrow aspirate suggested some dysplasticerythroid features, but a trephine biopsy revealed only a hypercellular marrow trephine with no hematological abnormalities; no evidence of hemophagocytic lymphohistiocytosis, tuberculosis nor leishmaniasis. She had been investigated for persistent diarrhea since the age of 1 year. Colonic biopsy at that time and again at presentation aged 19 demonstrated lymphocytic infiltrates associated with increased apoptosis in the colonic crypts. Crypt branching and lymphocytic infiltrates around the crypts were noted. These changes were consistent with a colitis (**Fig. 1d**), which had been managed conservatively. Her parents and siblings were, and remain, well (**Fig. 1a – left panels**). She had been suffering regular winter lower respiratory tract infections that resolved with antibiotic therapy. Despite having no persistent respiratory symptoms a high resolution CT scan of her chest showed nodular changes suggestive of a cellular infiltrate (**Fig. 1c – top right**). She received a course of high dose corticosteroids that immediately restored her platelet and neutrophil count and abrogated the pyrexia. Splenomegaly resolved 18 months after steroids, but colitis and radiographic lung changes still persist. Post-treatment, she remains lymphopenic and hypogammaglobulinemic **(Table 1** and **Supplementary Table 1)**. She has been unresponsive to pneumococcal or tetanus vaccinations. She has low B cell memory subsets and nearly absent B class switch recombination (**Supplementary Table 1**). This patient has been commenced on immunoglobulin therapy due to a progressive decline in IgG levels and increasing frequency of chest infections.

To identify a genetic defect in the patient, the whole exomes of all available family members were sequenced. As both siblings and parents were unaffected, the analysis focused on *de novo* and recessive modes of inheritance. After excluding all variants with minor allele frequency > 0.01 no candidate variants remained to support a hypothesis of recessive inheritance. One heterozygous variant appeared novel and *de novo.* The mutation was confirmed to be heterozygote in the patient and absent in family members using Sanger sequencing (**Fig. 1b – left and Supplementary Fig. 1**). This novel non-synonymous mutation in *BACH2,* c.T71C, leads to a leucine to proline substitution (p.L24P) (**Fig. 4a and Supplementary Table 2**). In the patient the variant was detected in blood and saliva samples suggesting that c.T71C mutation is germ-line rather than a somatic variant restricted to the bone marrow.

We identified a second family with similar clinical features and a heterozygous point mutation in *BACH2*, c.G2362A (causing p.E788K), found from whole exome sequencing in a father and daughter. In this second family (**Fig. 1a – right family**), a 64-year-old Caucasian male (subject B.II.1) presented at the age of 50 years with progressive shortness of breath associated with recurrent chest infections and sinusitis. At the age of 60 years he developed recurrent diarrhea. Investigations included CT of the chest that identified atelectasis and bronchiectasis together with mediastinal and hilar adenopathy (**Fig. 1c – lower panels**). He has low memory B cell subsets and profoundly reduced IgM, IgG and IgA levels (**Table 1** and **Supplementary Table 1**). He is currently receiving intravenous immunoglobulin (IvIg) therapy, which has had positive effect on sinusitis but the diarrhea persists and pulmonary symptoms have worsened over time. A daughter of the second patient (subject B.III.2) was diagnosed at the age of 10 years with ulcerative colitis. She underwent colectomy when 14 years old and has subsequently had recurrent pouchitis requiring antibiotics. When aged 32, her diagnosis was changed to Crohn's disease. She is currently 40 years old and remains troubled by recurrent attacks of IBD, lower and upper respiratory chest infections together with recurrent episodes of otitis media. She has low B cell memory subsets and undetectable serum IgA. Patient B.III.2 most likely falls into the selective IgA/CVID disease category given a family of antibody deficiency, typical parental offspring immunoglobulin, clinical history of recurrent sino-pulmonary infection, diagnosis of colitis and reduced total memory B cell profile[46-48]. She has been treated with TNFα blockers, but the treatment was discontinued because of alterations in kidney function. Sanger sequencing confirmed the c.G2362A mutation identified on whole exome sequencing in the two affected individuals and its absence in the healthy son (B.III.1) (**Supplementary Fig. 1**). c.G2362A leads to a glutamate to lysine substitution in the C-terminus of the protein (p.E788K) (**Fig. 4a and Supplementary Table 2)**. The clinical characteristics of all three patients are summarized in **Table 1** and **Supplementary Table 1.**