



HAL
open science

Identification of Germline Non-coding Deletions in XIAP Gene Causing XIAP Deficiency Reveals a Key Promoter Sequence

Zineb Sbihi, Kay Tanita, Camille Bachelet, Christine Bole, Fabienne Jabot-Hanin, Frederic Tores, Marc Le Loch, Radi Khodr, Akihiro Hoshino, Christelle Lenoir, et al.

► To cite this version:

Zineb Sbihi, Kay Tanita, Camille Bachelet, Christine Bole, Fabienne Jabot-Hanin, et al.. Identification of Germline Non-coding Deletions in XIAP Gene Causing XIAP Deficiency Reveals a Key Promoter Sequence. *Journal of Clinical Immunology*, 2022, 42 (3), pp.559-571. 10.1007/s10875-021-01188-z . hal-03864194

HAL Id: hal-03864194

<https://hal.science/hal-03864194>

Submitted on 23 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Identification of germline non-coding deletions in XIAP gene causing XIAP deficiency**
2 **reveals a key promoter sequence**

3 Zineb Sbihi¹, Kay Tanita², Camille Bachelet^{1,3}, Christine Bole⁴, Fabienne Jabot-Hanin^{4,5},
4 Frederic Tores^{4,5}, Marc Le Loch⁶, Radi Khodr¹, Akihiro Hoshino¹, Christelle Lenoir¹, Matias
5 Oleastro⁷, Mariana Villa⁷, Lucia Spossito⁷, Emma Prieto⁷, Silvia Danielian⁷, Erika Brunet⁸,
6 Capucine Picard^{1,3,9}, Takashi Taga¹⁰, Shimaa Said Mohamed Ali Abdrabou¹¹, Takeshi Isoda²,
7 Masafumi Yamada¹¹, Alejandro Palma⁷, Hiro Kanegane¹² and Sylvain Latour^{1,3}

8
9 ¹Laboratory of Lymphocyte Activation and Susceptibility to EBV infection, INSERM UMR
10 1163, Imagine Institute, Paris, France;

11 ²Department of Pediatrics and Developmental Biology, Graduate School of Medical and Dental
12 Sciences, Tokyo Medical and Dental University (TMDU), Tokyo, Japan;

13 ³Université de Paris, Paris, France;

14 ⁴Genomics Core Facility, Institut Imagine-Structure Fédérative de Recherche Necker, INSERM
15 U1163 et INSERM US24/CNRS UMS3633, Université de Paris, Paris, France;

16 ⁵Bioinformatic Platform, INSERM UMR 1163, Institut Imagine, Paris, France;

17 ⁶Service d'Histologie - Embryologie - Cytogénétique, Hôpital Necker-Enfants Malades, Paris,
18 France;

19 ⁷Immunology and Rheumatology Division, Hospital de Pediatria S.A.M.I.C. Prof. Dr. Juan P.
20 Garrahan, Buenos Aires, Argentina;

21 ⁸Laboratory of Dynamic of Genome and Immune System, INSERM UMR 1163, Imagine
22 Institute, Paris, France;

23 ⁹Study Center for Primary Immunodeficiencies, Necker-Enfants Malades Hospital, APHP,
24 Paris, France;

25 ¹⁰Department of Pediatrics, Shiga University of Medical Science, Otsu, Japan;

26 ¹¹Department of Pediatrics, Faculty of Medicine, Graduate School of Medicine, Hokkaido
27 University, Sapporo, Japan;

28 ¹²Department of Child Health and Development, Graduate School of Medical and Dental
29 Sciences, TMDU, Tokyo, Japan.

30
31 **Contact information:**

32 Corresponding author: Dr Sylvain Latour, E-mail: sylvain.latour@inserm.fr ; Tel: +33 (1) 42
33 75 43 03; Fax: +33 (1) 42 75 42 21.

34
35 Key words: X-linked inhibitor of apoptosis ; inherited immunodeficiency ; non-coding exon ;
36 deletion ; promoter

37 **Abstract**

38 **Purpose:** X-linked inhibitor of apoptosis protein (XIAP) deficiency, also known as the X-
39 linked lymphoproliferative syndrome of type 2 (XLP-2), is a rare immunodeficiency
40 characterized by recurrent hemophagocytic lymphohistiocytosis, splenomegaly and
41 inflammatory bowel disease. Variants in XIAP including missense, non-sense, frameshift and
42 deletions of coding exons have been reported to cause XIAP deficiency. We studied three young
43 boys with immunodeficiency displaying XLP-2-like clinical features. No genetic variation in
44 the coding exons of XIAP was identified by whole exome sequencing (WES), although the
45 patients exhibited a complete loss of XIAP expression.

46 **Methods:** Targeted next-generation sequencing (NGS) of the entire locus of XIAP was
47 performed on DNA samples from the three patients. Molecular investigations were assessed by
48 gene reporter expression assays in HEK cells and CRISPR-Cas9 genome editing in primary T
49 cells.

50 **Results:** NGS of XIAP identified three distinct non-coding deletions in the patients that were
51 predicted to be driven by repetitive DNA sequences. These deletions share a common region
52 of 839bp that encompassed the first non-coding exon of XIAP and contained regulatory
53 elements and marks specific of an active promoter. Moreover, we showed that among the
54 839bp, the exon was transcriptionally active. Finally, deletion of the exon by CRISPR-Cas9 in
55 primary cells reduced XIAP protein expression.

56 **Conclusions:** These results identify a key promoter sequence contained in the first non-coding
57 exon of XIAP. Importantly, this study highlights that sequencing of the non-coding exons that
58 are not currently captured by WES should be considered in the genetic diagnosis when no
59 variation is found in coding exons.

60 **Introduction**

61 X-linked inhibitor of apoptosis (XIAP) deficiency (OMIM #300635), also known as the
62 X-linked lymphoproliferative syndrome type 2 (XLP-2), is a recessive inherited primary
63 immunodeficiency estimated to affect approximately 1-3 males per million, although it may be
64 underdiagnosed. XIAP deficiency is associated with a high risk to develop inflammatory
65 symptoms including hemophagocytic lymphohistiocytosis (HLH) often in response to EBV
66 infection, recurrent splenomegaly, inflammatory bowel disease (IBD) with the features of
67 Crohn's disease [1-6]. Notably, heterozygous females can also develop symptoms as severe as
68 in males due to abnormal X-inactivation [4, 7].

69 XIAP was originally described as an anti-apoptotic molecule ubiquitously expressed. It
70 is composed of three Baculovirus inhibitor of apoptosis protein repeat (BIR) domains, a
71 ubiquitin binding domain (UBA), and a C-terminal RING domain with E3 ubiquitin ligase
72 activity. XIAP inhibits programmed cell death by binding to and blocking activated forms of
73 the effector caspases 3, 7 and 9 [8]. In addition to its anti-apoptotic function, XIAP is also
74 involved in several other pathways, including regulation and activation of innate immunity and
75 inflammation [9]. In particular, XIAP is required for signaling and activation of the nucleotide
76 oligomerization domain (NOD) receptors NOD1 and NOD2, which are intracellular sensors of
77 bacterial products [10, 11]. XIAP-deficient monocytes from patients display an impaired
78 production of cytokines and chemokines in response to stimulation with NOD2 ligands, while
79 XIAP-deficient T cells exhibit increased apoptosis when activated [1, 12, 13].

80 Mutations in XIAP gene have been reported in more than 100 patients worldwide since
81 2006 [1]. All the mutations described to date are in the coding exons and non-coding flanking
82 regions of exons of *XIAP*. Those include missense mutations and nonsense mutations (n=57),
83 mutations of canonical splice site (n=7), small insertions (n=10), small deletions (n=25), large

84 insertions/duplications (n=3) and deletions (n=15) (Web resources, The human Gene Mutation
85 Database, HGMD 2020.4) (**Figure S1**) [5-7, 11, 13-17].

86 To date, the contribution of disrupted potentially *cis*-regulatory conserved non-coding
87 sequences (CNCs) to human genetic diseases is largely underestimated, as no systematic screen
88 for putative deleterious variations in CNCs has been conducted. Several variants in non-coding
89 gene promoter and regulatory elements regions have been identified in genetic diseases. The
90 majority of them mediates their regulatory effects through alteration of transcriptional factors
91 binding and disturb trans-activation of the target gene promoter [18-21]. Promoter mutations
92 have also been identified in cancer predisposition genes, like in the alternative promoter of *APC*
93 that disrupts the binding of YY1 and reduces APC 1B promoter activity in human gastric cell
94 lines [22]. In familial melanoma, mutations in the TERT promoter altered the binding to
95 transcriptional factors leading to increased TERT promoter activity in human melanoma cell
96 lines [23]. Most recently, whole genome sequencing of a cohort of patients with sporadic
97 immunodeficiency identified compound heterozygosity of coding variants associated with non-
98 coding *cis*-regulatory element (CRE) deletions in *LRBA*, *DOCK8* and *ARPC1* genes,
99 demonstrating that sequence variations within non-coding regions of the genome have effects
100 on gene expression and can contribute to primary immunodeficiency [20]. Furthermore, recent
101 studies have provided insights into the biological relevance of the non-protein-coding portions
102 of the human genome [24]. One of them, the ENCODE pilot study revealed that functional
103 genomic elements are in much higher numbers than previously anticipated, and that the
104 majority of elements regulating gene expression are contained in the non-protein coding regions
105 [25].

106 In the present study, we report the identification of patients with XIAP deficiency whom
107 did not carry mutations in coding regions of *XIAP*, but distant deletions in non-protein coding

108 regions encompassing the non-coding exon 1 of *XIAP* that defines a key promoter sequence
109 required for XIAP expression.

110 **Methods**

111 *Exome sequencing and analysis.* Genomic DNA from whole blood or peripheral blood cells
112 of patients, their parents, and other family members was isolated according to standard
113 methods. Exome capture and analysis were performed as previously described [26].

114 *Targeted resequencing by NGS.* A custom capture by hybridization approach and method were
115 used as previously described [27-30]. The biotinylated single strand DNA probes targeting all
116 the region of interest were designed to cover a 143-kb chromosomal region including the *XIAP*
117 gene on chromosome X: 122904611-123047997 genomic position according to the
118 GRCh37/hg19 assembly of the human reference genome. The capture by hybridization libraries
119 were sequenced using a paired-end mode 100+100 bases on an Illumina NovaSeq6000, to
120 produce ~40 millions of clusters per sample. After demultiplexing, sequences were aligned to
121 the reference human genome hg19 using the Burrows-Wheeler Aligner [31]. The mean depth
122 of coverage per sample was ~200X. Downstream processing was carried out with the Genome
123 Analysis Toolkit (GATK), SAMtools and Picard, following documented best practices
124 (<http://www.broadinstitute.org/gatk/guide/topic?name=best-practices>). Variant calls were
125 made with the GATK Unified Genotyper. The annotation process is based on the latest release
126 of the Ensembl database. Variants were annotated and analyzed and prioritized using the
127 Polyweb/PolyDiag software interface designed by the Bioinformatics platform of University
128 Paris Descartes (Polyweb Imagine, <https://www.polyweb.fr>). Sequences were aligned with the
129 hg19 reference human genome with Burrows-Wheeler Aligner, version 0.6.2.13. CNVs were
130 analyzed using Integrative Genome Viewer (IGV) software
131 (<https://software.broadinstitute.org/software/igv/> or <https://igv.org>).

132 *Apoptosis assays.* $2 \cdot 10^5$ T-cell blasts/well were seeded into 96-well plates pre-coated with 0.1,
133 1 or 10 $\mu\text{g/ml}$ of mAb anti-CD3 OKT3 (eBioscience), or incubated with anti-CD3/CD28–
134 coated beads (Invitrogen). Apoptotic cells were detected after overnight stimulation by

135 propidium iodide staining according to standard protocol [1]. Cells were washed with PBS
136 twice, and data were acquired using a LSRFortessa (BD Biosciences) and analyzed using
137 FlowJo X (TreeStar). The percentage of induced apoptosis was calculated according to the
138 formula: $100 \times (\% \text{ of experimental apoptotic cells} - \% \text{ of spontaneously apoptotic cells} / 100 -$
139 $\% \text{ of spontaneously apoptotic cells})$.

140 ***Analysis of protein expression by immunoblotting.*** Protein extracts from PBMC and T-cell
141 blasts were obtained and analyzed as previously described [1, 26]. Proteins were separated by
142 SDS-PAGE and transferred on PVDF membranes (Millipore). Membranes were blocked with
143 milk for 1 hour before incubation with primary antibodies. The following antibodies were used
144 for immunoblotting: anti-XIAP (clone hILP, BD Bioscience), anti-KU80 (clone C48E7, Cell
145 Signaling) and anti-ACTIN (clone 13E5, Sigma-Aldrich). Membranes were then washed and
146 incubated with anti-mouse or anti-rabbit HRP-conjugated antibodies (Cell Signaling). Pierce
147 ECL western blotting substrate was used for detection (ThermoFisher). Densitometry analyses
148 were performed using ImageJ software and normalized to loading western blot KU80 controls.

149 ***In Silico analysis.*** To identify potential regulatory elements within the *XIAP* deletion region,
150 we used DNaseI sequencing data along with histone modification ChIP-seq data from
151 ENCODE (Encyclopedia of DNA elements) [32]. Bioinformatics analyses on deleted regions
152 were performed using publically available datasets from ENCODE, which includes information
153 such as the location of promoter and enhancer histone marks, open chromatin and bound
154 proteins.

155 ***Luciferase Assay.*** Primers were designed to amplify a genomic region that encompasses
156 potential *XIAP* regulatory regions. A 892bp product (chrX: 122993332–122994223)
157 corresponding to the commonly deleted region of *XIAP* sequence containing upstream
158 sequence, non-coding exon1, and partial intron 1 of the *XIAP* gene, or a 417bp product (chrX:
159 122993756–122994173) containing non-coding exon1, and partial intron1, or A 359bp product

160 (chrX: 122993421–122993779) containing upstream sequence of the *XIAP* gene, were
161 amplified from control genomic DNA, cloned into the promoter-less firefly luciferase reporter
162 vector pGL3-Basic (Promega). Primers used for cloning incorporated KpnI and XhoI restriction
163 site to facilitate subcloning are listed in **Table S1**. All the constructs were verified by Sanger
164 sequencing (Applied Biosystems). pGL3-Basic plasmids containing the promoter sequences of
165 *XIAP* (90ng) were co-transfected with 10 ng of pRL-CMV (CMV-promoter driven *Renilla*
166 luciferase reporter, Promega) in HEK293 cells using lipofectamine 2000 (Invitrogen) according
167 to the manufacturer's instructions. 24h post-transfection, luciferase activity was measured using
168 a dual-luciferase reporter assay system according to the manufacturer's instructions (dual-Glo
169 Luciferase reporter assay, Promega).

170 ***Cas9 RNP-mediated Editing of primary human T cells.*** T-cell blasts were activated with anti-
171 CD3/CD28-coated beads (Invitrogen) for 48 hours prior to electroporation with
172 ribonucleoprotein/RNP complexes using the P3 Primary cell 4D X kit and 4D-Nucleofector
173 (Lonza). The *Streptococcus pyogenes* Cas9 protein with 2 nuclear localization signals (NLS)
174 expressed and purified as described by Menoret et al. [33], was kindly provided by Dr. J.P.
175 Concordet (Structure and Instability of Genomes laboratory, Muséum National d'Histoire
176 Naturelle, France). The following *XIAP* gene-specific sgRNAs, designed using the online tool:
177 <http://crispor.tefor.net> [34], were used: sgRNA1: 5'-CUGGGAUAGUUAUCCCCUGU-3' ;
178 sgRNA2: 5'-UUCCUCGGACUGCCGACGGC-3' ; sgRNA3: 5'-
179 CCAGCCCGGGCUGCGCCACU-3'; sgRNA4: 5'-GCGGUGGGUACAGCUUGUGU-3'.
180 Chemically modified sgRNAs were purchased from Synthego with protospacer adjacent motif
181 (<https://www.synthego.com>). To generate Cas9 RNP complexes, *Streptococcus pyogenes* Cas9
182 protein (90µM) and sgRNAs (180µM) were mixed 1 : 2. For each condition, 4.10⁶ stimulated
183 T-cell blasts were transfected with the Cas9 RNP mix using program CA-137 on the Amaxa
184 4D-Nucleofector (Lonza). 300µL pre-warmed complete medium was added to each well, and

185 the cells were allowed to recover for 10 min at room temperature. Cells were then plated in 1
186 mL of medium with 100IU of IL2, and maintained at 1.10^6 cells / mL for 20 days. The region
187 on genomic DNA that spans the cutting site of each sgRNA (sgRNA1, sgRNA2, sgRNA3, and
188 sgRNA4) was amplified by PCR using the following on-target primers: forward: 5'-
189 CTTTGTTCCTCCGGTCCATCTGC-3' and reverse: 5'-CGGAAGCTCTTTGGCCCTTA-3';
190 forward: 5'-AAGAAACACTGGAGCTGGGG-3' and reverse: 5'-
191 AATCCTGCAGGCCTGAAGTC-3', respectively (Eurofin). Amplified products were
192 separated by agarose gel electrophoresis and Sanger sequenced (Applied-Biosystems).

193 **Results**

194 **Identification of three unrelated patients presenting XIAP deficiency-like features**

195 We investigated three male patients with immunodeficiency from three unrelated
196 families one from Argentina (Patient 1, P1), and two from Japan (Patient 2, P2 and Patient 3,
197 P3) (**Figure 1A**). P1 became symptomatic within the first months of life, P2 by the age of 18
198 months, and P3 by the age of 9 years. All patients presented with recurrent **severe** HLH
199 triggered (P1) or not (P2 and P3) by EBV infection; this was the first clinical manifestation in
200 the three patients (**Table 1**). In P1 and P2, HLH was associated with persistent splenomegaly
201 associated with fluctuating thrombocytopenia in P1 and pancytopenia in P2. P1 was treated by
202 cyclosporine and prednisone. P2 was treated by dexamethasone and was refractory to
203 intravenous immunoglobulin. P3 developed inflammatory bowel disease (IBD) concomitantly
204 to HLH and was refractory to treatments with steroid, infliximab and tacrolimus. He underwent
205 several colectomies between 9 and 18 years old, and finally received allogeneic hematopoietic
206 stem cell transplantation at 19 years of age and he is well since (**Table 1**).

207 **Absence of XIAP expression in cells of patients**

208 Because HLH, IBD and splenomegaly are clinical features reported to be associated
209 with XIAP deficiency [35], XIAP deficiency was evaluated as a possible cause of the clinical
210 symptoms in the three patients. XIAP expression was first examined in PBMCs from P1 and
211 P2 by intracellular staining by flow cytometry suggesting a reduced XIAP expression (**Figure**
212 **S2A and data not shown**). However, we experienced that this method frequently led to false
213 positive as exemplified by a patient with a deletion in the coding exon 3 of XIAP with apparent
214 normal XIAP expression by flow cytometry, while there was no XIAP expression by western
215 blot (**Figure S3 and data not shown**). Thus, XIAP expression was assessed by western blot in
216 cell extracts of T lymphocytes from P1, P2 and P3. No detectable expression was observed in
217 cells of P1, P2 and P3, in contrast to the reliable XIAP expression detected in lysates of control

218 T cells from healthy donors (**Figure 1B and data not shown for P3**). Moreover, analysis of
219 *XIAP* transcripts by PCR in cells of patients P1, P2 and P3 failed to amplify *XIAP* transcripts
220 in comparison to the control cells in which *XIAP* transcripts were detected (**Figure 1C and**
221 **data not shown for P3**). Increased activation-induced cell-death of T lymphocytes and
222 defective NOD2-mediated IL-8 production by myeloid cells, which represent two major
223 hallmarks of the *XIAP* deficiency [1, 4], were observed respectively in T cells of patient P1 and
224 in PBMCs of P2 and P3 (**Figure S2B and C**). Therefore, these data strongly suggested that the
225 three patients had a *XIAP* deficiency explaining their clinical phenotypes.

226 **Identification of large deletions in 5' non-coding regions of *XIAP* in three patients**

227 We thus investigated the *XIAP* gene in the three patients for genetic variations that could
228 explain their loss of *XIAP* expression. Surprisingly, whole-exome sequencing (WES)
229 performed in P1 failed to identify genetic variations in *XIAP*, although the sequencing coverage
230 of the six coding exons of *XIAP* was significant (**data not shown**). In patients P2 and P3,
231 targeted Sanger sequencing of the 6 coding exons including non-coding flanking regions (of
232 50-80bp) was conducted and did not reveal any genetic variations (**data not shown**). Therefore,
233 in P1, P2 and P3, the *XIAP* deficiency was not caused by mutations in 6 coding exons of *XIAP*,
234 suggesting another mechanism(s), since all mutations in *XIAP* reported so far to cause *XIAP*
235 deficiency target the 6 coding exons (**Figure 1S**).

236 Given that variants within a non-coding or regulatory region are missed by WES or by
237 targeted sequencing, we next performed high-throughput targeted resequencing with custom
238 capture oligonucleotides designed to cover a large region of 87386bp encompassing the *XIAP*
239 locus (from g.122960611 to g.123047997) (**Figure 2**). This region includes 80 kb of sequences
240 5' to the start codon and 7 kb of the sequences 3' to the stop codon. Patients and their parents
241 (when available) as well as two healthy controls were analyzed. We searched for rare single-
242 nucleotide variants (SNVs), insertions and deletions (indels). Confirming the first analyses, no

243 SNVs or indels in the 6 coding exons were detected in the three patients. However, there was
244 no sequence reads in regions between the 5' region and the first intron 1 encompassing the non-
245 coding exon1 in P1, P2 and P3, indicating hemizygous deletions in the 5' region of *XIAP* in the
246 three patients. In P2, one short deletion and one long deletion were detected. The coverage (or
247 numbers of reads) of this region in the mother of P1 was twice that in the father and in a healthy
248 male control, suggesting that the hemizygous deletion in P1 was *de novo* or transmitted through
249 germline mosaicism. The coverage of the long deletion in the mother of P2 was equivalent to a
250 healthy male control and reduced about twice compared to a healthy female control and around
251 50% of reads carried the same deleted region as in P2, suggesting that the hemizygous long
252 deletion in P2 was inherited from the mother (**Figure 2** and **Figure S4**). Similarly, the short
253 deletion in P2 was also inherited from the mother, but was found to be homozygous in the
254 mother as no wild-type reads were detected (**Figure S4**). In family 3, the mother of P3 was not
255 analyzed by NGS, but the deleted region was successfully amplified by PCR from DNA of the
256 mother strongly supporting mother inheritance of the deletion (**Figure S4C**).

257 To confirm these findings and precisely map the limits of these deletions, targeted
258 amplification by PCR and Sanger sequencing of surrounding regions of deletions were
259 performed. In P1 and P3, amplicons were obtained and sequenced, demonstrating hemizygous
260 deletions of 21,521bp (chrX: g.122979032_123000553del, hereafter denoted as $\Delta 1$) and
261 2,199bp (chrX: g.122993375_122995574del, hereafter denoted as $\Delta 3$), respectively (**Figure**
262 **3A and B and Figure S4**). PCR products were amplified in mother of P3, but not in mother of
263 P1 confirming the maternal transmission in P3 and the *de novo* or maternal transmission
264 through germline mosaicism in P1 (**Figure S4**). In P2, we failed to amplify the regions
265 surrounding the two deletions in P2 possibly due of the GC enrichment. Hence, the limits of
266 the deletions in P2 were only assessed by analysis of NGS data showing two hemizygous
267 deletions of 862bp and 17,979bp (chrX: g.122972822_122973684del and

268 g.122976235_122994214del, hereafter denoted $\Delta 2a$ and $\Delta 2b$) (**Figure S5 and Figure 3A and**
269 **3B**). Therefore, these data identified deletions in 5' non-coding regions in *XIAP* in the three
270 patients that likely account for their *XIAP* deficiency.

271 **Analysis of repeat elements at the *XIAP* locus and mechanisms underlying the three** 272 **deletions**

273 Double strand breaks (DSBs) when occurring in repeated elements are associated with
274 genome instability including deletions. DSB repair by homologous recombination or by a non-
275 homologous end-joining repair mechanism may lead to crossovers between two repeated
276 sequences or micro-homologies in sequences resulting in the excision of the intervening region
277 [36, 37]. Thus, we next investigated whether the breakpoint regions of the three deletions were
278 located within repeated elements (**Figure 3A**). First, using RepeatMasker program (see Web
279 resources), we determined the percentage of repeated elements including *Alu* elements, *Line*
280 elements, DNA transposons and C-rich elements in the *XIAP* locus. We found that 64.58% of
281 the *XIAP* locus consists of repeated elements including 42.97% of *Alu* elements, 18.17% of
282 long Interspersed Nuclear (*LINEs*) elements and 1.9% of simple sequence repeats. Repeated
283 elements represent 62% of the genome of the X chromosome and among them 10% are *Alu*
284 elements (from Bioinfo-fr, Human genome repeated elements, see Web resources). The *XIAP*
285 locus can thus be considered to be enriched in repeated elements in particular *Alu* elements. As
286 expected these repeated elements are preferentially located in the non-coding regions of the
287 *XIAP* locus: the 5' region, 3' region, and intronic regions (**Figure 3A**). Alignments of the
288 breakpoints with UCSC Genome Browser data [38] showed that the four deletions were located
289 within 4 different repeat elements: *AluSx* in the 5' region and *AluSx4* in intron1-2 for $\Delta 1$; *AluSq2*
290 in the 5' region and *AluSx1* elements in intron 1-2 for $\Delta 2a$; *AluSp* in the 5' region and *C-rich*
291 elements in intron 1-2 for $\Delta 2b$; and one *MIR* in the 5' region and low complexity repeats
292 consisting of DNA transposons n(TA) in intron 1-2 for $\Delta 3$ (**Figure 3B and Table S2**).

293 Junctional micro-homologies observed at the breakpoints suggested that deletions were formed
294 by DSBs repair mechanisms such as replicative based-repair, non-homologous end-
295 joining/NHEJ and/or alternative end-joining (also termed microhomology mediated end-
296 joining/MMEJ) mechanisms [36]. Another possible mechanism is the non-allelic homologous
297 recombination that could occur between repetitive elements of the same family such in $\Delta 1$ and
298 $\Delta 2a$. Microhomology of 1-4bp rather facilitate NHEJ that may account for $\Delta 1$ and $\Delta 3$, while
299 microhomology up to 5bp as found in $\Delta 2a$ and $\Delta 2b$ may favor a MMEJ-dependent mechanism.
300 The deletion of one nucleotide (C) in $\Delta 3$ at the breakpoints strongly supports a NHEJ or MMEJ-
301 dependent repair mechanism at work, as both are prone to errors and could lead to base excision.
302 However, these observations indicated that the three deletions are located in repeated sequences
303 prone to genomic rearrangements.

304 **The first non-coding exon in *XIAP* deleted in the three patients hold characteristics of a**
305 **promoter sequence**

306 Importantly, intersection of the three deletions revealed a commonly deleted region of
307 839bp in the three patients (g.122993375_122994214del). This deleted region contains 5' UTR
308 sequences and the first non-coding exon1 of *XIAP* of 267 pb (chrX: g.122993877_122994143)
309 (**Figures 2 and 4 and Figure S4**). Using ENCODE databases, we analyzed this region to
310 determine whether it contains regulatory and/or promoter elements (**Figure 4A and Figure**
311 **S6**). The 839bp region was predicted to be enriched in regulatory promoter elements
312 corresponding to a putative promoter (between chrX: g.122992320 and 122996142
313 corresponding the red box in Figure 3A), **Figure 4A**). Notably, interrogation of chromatin
314 immunoprecipitation sequencing (ChIP-seq) ENCODE data revealed that this region is
315 transcriptionally active with multiple binding sites for different transcription factors (**Figure**
316 **4A and Table S3**). Moreover, this region was found within an accessible and active chromatin
317 region as determined by ENCODE derived DNase I hypersensitive clusters and enrichment for

318 promoter histone modifications (H3K4me3) and active histone modifications (H3K27ac).
319 These observations strongly suggested that this region has a promoter activity and, thus its
320 deletion accounts for the XIAP deficiency in the patients.

321 **The first non-coding exon in *XIAP* contains active promoter sequences required for XIAP**
322 **expression**

323 To further characterize the role of the non-coding exon1 as a promoter sequence, the
324 839bp region was divided in two, one containing 3'sequences corresponding to exon1 and one
325 corresponding to 5'sequences flanking exon1 (**Figure 4B**). The entire 839bp and the two sub-
326 regions were cloned in a gene expression reporter vector at the 5' of the ATG of the luciferase
327 gene. The vectors were transiently transfected in HEK cells and luciferase activity was
328 measured (**Figure 4C**). Luciferase activity was significantly detectable in cell-extracts of cells
329 transfected with the entire 839bp region or the 3'sequences containing exon1, but not in extracts
330 from cells transfected with the 5'sequences flanking exon1. As a positive control of luciferase
331 activity, cells were transfected with a vector coding the transcription factor EOMES in
332 combination with the reporter gene vector, in which the luciferase was under the control of the
333 proximal *IFNG* promoter that is known to be activated by EOMES. Therefore, these data
334 indicate that the 839bp region has a promoter activity that is contained in sequences containing
335 the non-coding exon1.

336 To formally demonstrate that the non-coding exon1 of *XIAP* contained a key promoter
337 sequence required for XIAP expression, we used CRISPR/Cas9 gene editing to obtain primary
338 T cells from healthy control males in which the non-coding exon1 was deleted (**Figure 5**). We
339 designed two pairs of single-guide RNAs (sgRNAs), one pair targeting a locus containing the
340 839bp region commonly deleted in the three patients, sgRNA1 and sgRNA2, while the second
341 targeted a locus restricted to *XIAP* non-coding exon1. Cleavages by sgRNA1 and sgRNA2 or
342 by sgRNA3 and sgRNA4 theoretically resulted respectively in a 1050bp deletion

343 (XIAP Δ 1050bp, containing the 839bp region) and in a 300bp deletion (XIAP Δ 300bp,
344 containing the exon1) (**Figure 5A**). XIAP-targeting sgRNA guides complexed with the Cas9
345 protein were delivered to activated T cell blasts, and cells were further expanded for genetic,
346 molecular and functional validation of the gene editing. Two regions of 1456bp and 843bp
347 encompassing the 1050bp and 300bp deletions respectively were amplified from DNA
348 extracted from cells following expansion at day 5 and day 13 after Cas9-sgRNAs delivery. Only
349 PCR products from cells that received Cas9 complexed with both sgRNA1 and sgRNA2 or
350 both sgRNA3 and sgRNA4 showed specific deletions of 1050bp and 300bp respectively
351 (**Figure 5B**). However, wild-type sequences were still present indicating that only a fraction of
352 cells has been targeted. **Since we used pairs of sgRNAs to create deletions, we could not**
353 **evaluate the % of edited cells using TIDE or ICE tools. However, based on the intensities of**
354 **the PCR products in the gel, we estimated that roughly 30-40% of cells carried the 1050bp**
355 **deletion and around 50% of cells had the 300pb deletion.** In contrast, no specific deletion was
356 observed with Cas9-RNPs containing single sgRNA, scramble sgRNA or no sgRNA. The
357 deletions were confirmed by Sanger sequencing of the PCR products (**Figure 5C**).
358 Immunoblotting of lysates from cells having received Cas9-RNPs containing both sgRNA1 and
359 sgRNA2 or both sgRNA3 and sgRNA4 showed a significant reduction of XIAP expression
360 relative to non-targeting controls (**Figure 5D**). The expression of XIAP was not completely
361 abolished likely because only a fraction of cells was targeted as indicated by the PCR analyses
362 (**Figure 5B**). Finally, this reduction in XIAP expression was associated with an increase of
363 activation-induced cell death/apoptosis in T cells stimulated with anti-CD3 antibody (**Figure**
364 **5E**). Taken together, these data demonstrated that the non-coding exon1 contains key promoter
365 sequences required for XIAP protein expression.

366

367 **Discussion**

368 The genetic basis of diseases is mainly focused on coding regions. However, more and
369 more evidence is coming now, that a large proportion of non-coding sequences in the genome
370 is functional and harbors genetic variants which contribute to disease etiology. In this study, we
371 reported three XIAP-deficient young male patients from three unrelated kindreds originating
372 from Japan and Argentina. The three patients studied have XLP-2-like clinical features,
373 associated with typical manifestation with HLH, IBD and splenomegaly. Surprisingly, no
374 genetic variation in the coding exons of XIAP was identified by whole exome sequencing
375 (WES), although the patients exhibited a complete loss of XIAP expression. We reported here
376 first XIAP-deficiency associated with a large non-coding hemizygous deletions. The deletions
377 reported here lead to a complete absence of XIAP expression in patient cells. The analysis of
378 the breakpoints of the three deletions described in the study, showed that the deletions involved
379 Alu elements, *C-rich* elements, MIR and low complexity repeats consisting of DNA
380 transposons n(TA), indicating that the three deletions are located in repeated sequences prone
381 to genomic rearrangements.

382 The first non-coding exon of XIAP was commonly removed in these three deletions, we
383 showed that this region has a promoter activity, and by CRISPR-Cas9 genome editing in
384 primary T cells we demonstrated that the non-coding exon1 contains key promoter sequences
385 required for XIAP protein expression. Importantly, non-coding exons are not captured by WES
386 and are often not included in targeted sequencing panels for different genetic diseases including
387 those for primary immunodeficiencies (PIDs). We further analyzed 463 genes known to cause
388 PIDs for the presence of non-coding exons. 42% (196 genes) of the genes harbors non-coding
389 exons, and among them 79% (156 genes) are predicted to contain regulatory promoter elements
390 based on ENCODE (representing a total of 40609bp) (**data not shown**). This is thus rather
391 surprising that non-coding exons are still not well considered in the different genetic diagnostic

392 tools. Importantly, our results highlight and provide a clear demonstration that genetic
393 variations in non-coding exons can lead to genetic disease and should be considered for
394 sequencing when no genetic variations in coding regions are identified.

395 Our study also points out that protein expression analysis, that can be viewed as time
396 consuming, remains one accurate way to achieve an unambiguous diagnosis, when in first
397 intention genetic analysis fails.

398

399

400

401

402 **Supplementary Information**

403 The online version contains supplemental data that include supplemental material and methods,
404 6 supplemental figures and 3 tables.

405 **Web resources**

406 Online Medelian Inheritance in Man, <http://www.omim.org/>

407 ENCODE data in the UCSC Genome Browser, <https://genome.ucsc.edu/>

408 RepeatMasker program, software package: Smit A., (RepeatMasker Open, 2013-
409 2015), <https://www.repeatmasker.org/faq.html/>

410 Bioinfo-fr, Human genome repeated elements, Devailly G., (Bioinfo-fr, 2017), [https://bioinfo-](https://bioinfo-fr.net/)
411 [fr.net/](https://bioinfo-fr.net/)

412 Human Gene Mutation Database, HGMD 2020.4, Human Gene Mutation Database - Cardiff
413 University, <http://www.hgmd.cf.ac.uk>

414 **Funding**

415 This work was supported by grants from the Ligue Contre le Cancer-Equipe Labélisée (France;
416 to S. L.), Institut National de la Santé et de la Recherche Médicale (France), exome sequencing
417 was funded by the Rare Diseases Foundation (France; to S.L.), the Agence Nationale de
418 Recherche (ANR, France) (ANR-18-CE15-0025-01 to S.L. and ANR-10-IAHU-01 to Institut
419 Imagine), the Société Française de Lutte contre les Cancers et Leucémies de l'Enfant et de
420 l'Adolescent, AREMIG (France; to S.L.), and the Fédération Enfants et Santé (France ; to S.L.).
421 S. L. is a senior scientist at the Centre National de la Recherche Scientifique (France). Z.S is
422 supported by the Fondation ARC pour la recherche sur le Cancer France.

423 **Conflict of interest/Competing interests**

424 The authors declare no potential conflicts of interest.

425 **Availability of data and material**

426 All material and data are available on request

427 **Code availability**

428 All code are available on request

429 **Author's contributions**

430 Z.S., K.T. and C.B. designed, performed experiments and analyzed the data. K.T. A.H., F. J-H,

431 F.T., M.L.L., C.B. and C.L. performed experiments and analyzed the data. A.H., H.K., A.P.,

432 T.I., S.S, T.T. M.Y. identified the patients, provided clinical and analyzed the data. S.L. and

433 Z.S. wrote the manuscript. S. L. designed and supervised the research.

434 **Ethics approval**

435 The study and protocols conform to the 1975 Declaration of Helsinki as well as to local

436 legislation and ethical guidelines from the Comité de protection des personnes de l'Ile de France

437 II and the French advisory committee on data processing in medical research.

438 **Consent to participate**

439 Informed and written consent was obtained from donors, patients and families of patients.

440 **Consent for publication**

441 All authors consent for publication of the manuscript

442

443 **References**

444

- 445 1. Rigaud S, Fondaneche MC, Lambert N, Pasquier B, Mateo V, Soulas P, et al. XIAP
446 deficiency in humans causes an X-linked lymphoproliferative syndrome. *Nature*.
447 2006;444(7115):110-4.
- 448 2. Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, Decker B, et al.
449 Making a definitive diagnosis: successful clinical application of whole exome sequencing in a
450 child with intractable inflammatory bowel disease. *Genet Med*. 2011;13(3):255-62.
- 451 3. Speckmann C, Ehl S. XIAP deficiency is a mendelian cause of late-onset IBD. *Gut*.
452 2014;63(6):1031-2.
- 453 4. Aguilar C, Lenoir C, Lambert N, Begue B, Brousse N, Canioni D, et al.
454 Characterization of Crohn disease in X-linked inhibitor of apoptosis-deficient male patients
455 and female symptomatic carriers. *J Allergy Clin Immunol*. 2014;134(5):1131-41 e9.
- 456 5. Ono S, Okano T, Hoshino A, Yanagimachi M, Hamamoto K, Nakazawa Y, et al.
457 Hematopoietic Stem Cell Transplantation for XIAP Deficiency in Japan. *J Clin Immunol*.
458 2017;37(1):85-91.
- 459 6. Chen RY, Li XZ, Lin Q, Zhu Y, Shen YY, Xu QY, et al. Epstein-Barr virus-related
460 hemophagocytic lymphohistiocytosis complicated with coronary artery dilation and acute
461 renal injury in a boy with a novel X-linked inhibitor of apoptosis protein (XIAP) variant: a
462 case report. *BMC Pediatr*. 2020;20(1):456.
- 463 7. Yang X, Hoshino A, Taga T, Kunitsu T, Ikeda Y, Yasumi T, et al. A female patient
464 with incomplete hemophagocytic lymphohistiocytosis caused by a heterozygous XIAP
465 mutation associated with non-random X-chromosome inactivation skewed towards the wild-
466 type XIAP allele. *J Clin Immunol*. 2015;35(3):244-8.
- 467 8. Eckelman BP, Salvesen GS, Scott FL. Human inhibitor of apoptosis proteins: why
468 XIAP is the black sheep of the family. *EMBO Rep*. 2006;7(10):988-94.
- 469 9. Yabal M, Muller N, Adler H, Knies N, Gross CJ, Damgaard RB, et al. XIAP restricts
470 TNF- and RIP3-dependent cell death and inflammasome activation. *Cell Rep*.
471 2014;7(6):1796-808.
- 472 10. Krieg A, Correa RG, Garrison JB, Le Negrato G, Welsh K, Huang Z, et al. XIAP
473 mediates NOD signaling via interaction with RIP2. *Proc Natl Acad Sci U S A*.
474 2009;106(34):14524-9.
- 475 11. Parackova Z, Milota T, Vrabcova P, Smetanova J, Svaton M, Freiburger T, et al.
476 Novel XIAP mutation causing enhanced spontaneous apoptosis and disturbed NOD2
477 signalling in a patient with atypical adult-onset Crohn's disease. *Cell Death Dis*.
478 2020;11(6):430.
- 479 12. Damgaard RB, Nachbur U, Yabal M, Wong WW, Fiil BK, Kastirr M, et al. The
480 ubiquitin ligase XIAP recruits LUBAC for NOD2 signaling in inflammation and innate
481 immunity. *Mol Cell*. 2012;46(6):746-58.
- 482 13. Damgaard RB, Fiil BK, Speckmann C, Yabal M, zur Stadt U, Bekker-Jensen S, et al.
483 Disease-causing mutations in the XIAP BIR2 domain impair NOD2-dependent immune
484 signalling. *EMBO Mol Med*. 2013;5(8):1278-95.
- 485 14. Marsh RA, Madden L, Kitchen BJ, Mody R, McClimon B, Jordan MB, et al. XIAP
486 deficiency: a unique primary immunodeficiency best classified as X-linked familial
487 hemophagocytic lymphohistiocytosis and not as X-linked lymphoproliferative disease. *Blood*.
488 2010;116(7):1079-82.
- 489 15. Filipovich AH, Zhang K, Snow AL, Marsh RA. X-linked lymphoproliferative
490 syndromes: brothers or distant cousins? *Blood*. 2010;116(18):3398-408.
- 491 16. Zeissig Y, Petersen BS, Milutinovic S, Bosse E, Mayr G, Peuker K, et al. XIAP
492 variants in male Crohn's disease. *Gut*. 2015;64(1):66-76.

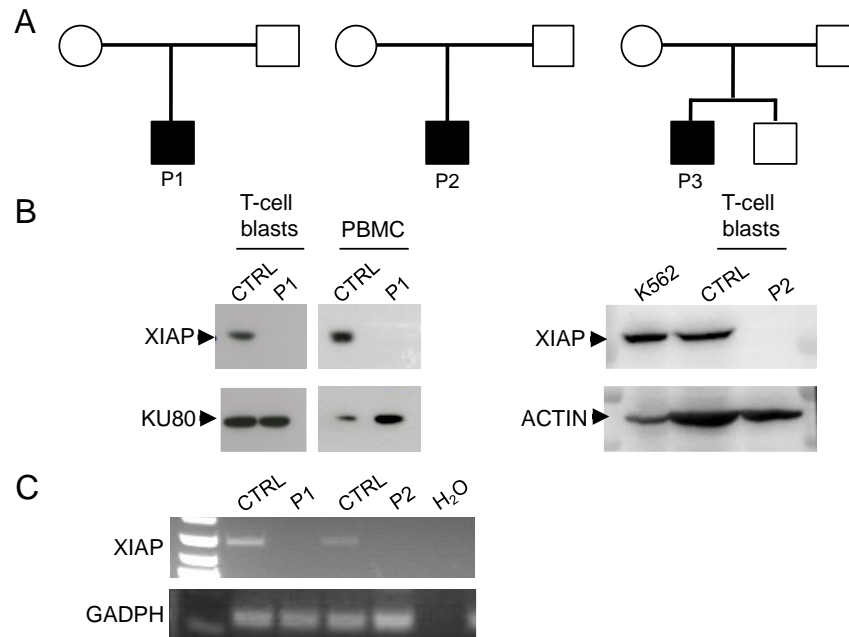
- 493 17. Latour S, Aguilar C. XIAP deficiency syndrome in humans. *Semin Cell Dev Biol.*
494 2015;39:115-23.
- 495 18. Davidson AE, Liskova P, Evans CJ, Dudakova L, Noskova L, Pontikos N, et al.
496 Autosomal-Dominant Corneal Endothelial Dystrophies CHED1 and PPCD1 Are Allelic
497 Disorders Caused by Non-coding Mutations in the Promoter of OVOL2. *Am J Hum Genet.*
498 2016;98(1):75-89.
- 499 19. Wakabayashi A, Ulirsch JC, Ludwig LS, Fiorini C, Yasuda M, Choudhuri A, et al.
500 Insight into GATA1 transcriptional activity through interrogation of cis elements disrupted in
501 human erythroid disorders. *Proc Natl Acad Sci U S A.* 2016;113(16):4434-9.
- 502 20. Thaventhiran JED, Lango Allen H, Burren OS, Rae W, Greene D, Staples E, et al.
503 Whole-genome sequencing of a sporadic primary immunodeficiency cohort. *Nature.*
504 2020;583(7814):90-5.
- 505 21. Turro E, Astle WJ, Megy K, Graf S, Greene D, Shamardina O, et al. Whole-genome
506 sequencing of patients with rare diseases in a national health system. *Nature.*
507 2020;583(7814):96-102.
- 508 22. Li J, Woods SL, Healey S, Beesley J, Chen X, Lee JS, et al. Point Mutations in Exon
509 1B of APC Reveal Gastric Adenocarcinoma and Proximal Polyposis of the Stomach as a
510 Familial Adenomatous Polyposis Variant. *Am J Hum Genet.* 2016;98(5):830-42.
- 511 23. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, et al. TERT promoter
512 mutations in familial and sporadic melanoma. *Science.* 2013;339(6122):959-61.
- 513 24. French JD, Edwards SL. The Role of Noncoding Variants in Heritable Disease. *Trends*
514 *Genet.* 2020;36(11):880-91.
- 515 25. Gerstein MB, Kundaje A, Hariharan M, Landt SG, Yan KK, Cheng C, et al.
516 Architecture of the human regulatory network derived from ENCODE data. *Nature.*
517 2012;489(7414):91-100.
- 518 26. Martin E, Palmic N, Sanquer S, Lenoir C, Hauck F, Mongellaz C, et al. CTP synthase
519 1 deficiency in humans reveals its central role in lymphocyte proliferation. *Nature.*
520 2014;510(7504):288-92.
- 521 27. Benyelles M, Episkopou H, O'Donohue MF, Kermasson L, Frange P, Poulain F, et al.
522 Impaired telomere integrity and rRNA biogenesis in PARN-deficient patients and knock-out
523 models. *EMBO Mol Med.* 2019;11(7):e10201.
- 524 28. Venot Q, Blanc T, Rabia SH, Berteloot L, Ladraa S, Duong JP, et al. Targeted therapy
525 in patients with PIK3CA-related overgrowth syndrome. *Nature.* 2018;558(7711):540-6.
- 526 29. Rosain J, Oleaga-Quintas C, Deswarte C, Verdin H, Marot S, Syridou G, et al. A
527 Variety of Alu-Mediated Copy Number Variations Can Underlie IL-12Rbeta1 Deficiency. *J*
528 *Clin Immunol.* 2018;38(5):617-27.
- 529 30. Simonetti L, Bruque CD, Fernandez CS, Benavides-Mori B, Delea M, Kolomenski JE,
530 et al. CYP21A2 mutation update: Comprehensive analysis of databases and published genetic
531 variants. *Hum Mutat.* 2018;39(1):5-22.
- 532 31. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler
533 transform. *Bioinformatics.* 2010;26(5):589-95.
- 534 32. Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR,
535 et al. Identification and analysis of functional elements in 1% of the human genome by the
536 ENCODE pilot project. *Nature.* 2007;447(7146):799-816.
- 537 33. Menoret S, De Cian A, Tesson L, Remy S, Usal C, Boule JB, et al. Homology-
538 directed repair in rodent zygotes using Cas9 and TALEN engineered proteins. *Sci Rep.*
539 2015;5:14410.
- 540 34. Concordet JP, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9
541 genome editing experiments and screens. *Nucleic Acids Res.* 2018;46(W1):W242-W5.

- 542 35. Aguilar C, Latour S. X-linked inhibitor of apoptosis protein deficiency: more than an
543 X-linked lymphoproliferative syndrome. *J Clin Immunol.* 2015;35(4):331-8.
- 544 36. van Zelm MC, Geertsema C, Nieuwenhuis N, de Ridder D, Conley ME, Schiff C, et
545 al. Gross deletions involving IGHM, BTK, or Artemis: a model for genomic lesions mediated
546 by transposable elements. *Am J Hum Genet.* 2008;82(2):320-32.
- 547 37. Verdin H, D'Haene B, Beysen D, Novikova Y, Menten B, Sante T, et al.
548 Microhomology-mediated mechanisms underlie non-recurrent disease-causing microdeletions
549 of the FOXL2 gene or its regulatory domain. *PLoS Genet.* 2013;9(3):e1003358.
- 550 38. Casper J, Zweig AS, Villarreal C, Tyner C, Speir ML, Rosenbloom KR, et al. The
551 UCSC Genome Browser database: 2018 update. *Nucleic Acids Res.* 2018;46(D1):D762-D9.
552
553

Table 1. Clinical presentation and data of patients.

	Patient 1	Patient 2	Patient 3
Age at initial presentation	1 m. o.	1 y. 6 m. o.	1 y. o. HLH ; 9 y.o. colitis
Current age	9 y.	5 y.	21 y.
Origin	South American (Argentina)	Asian (Japan)	Asian (Japan)
Family history	-	-	-
HLH	+	+	+
Recurrent HLH	+	+	+
Fever	+	-	-
Splenomegaly	+ (persistent, never resolved)	-	-
Cytopenia	+	+	-
Thrombocytopenia	+	+	-
EBV	+	-	-
Other infections	Rotavirus diarrhea and UTI by K. pneumoniae at 2 m. o. ; CMV reactivation during admission for HLH at 2 y. o. ; 1 episode of pneumonia at 5 y. o. (treated with amoxicilin-clavulanate)	-	-
Hypogammaglobunemia	-	-	-
Inflammatory bowel disease/colitis	-	-	+
Treatments	Methylprednisolone, + anti-thymocyte globulin (for HLH) ; Rituximab (for HLH) ; maintenance therapy with cyclosporine and SCIG	Dexamethasone (for HLH)	Methylprednisolone (for HLH), Prednisolone, Infliximab, Tacrolimus, Adalimumab (for colitis)
Allogeneic HSCT	+	-	+
Surgery	-	-	Hemicolectomy, Total colectomy, Ileostomy

556 Abbreviations : HLH, hemophagocytic lymphohistiocytosis ; ND, no data ; EBV, Epstein-Barr virus ; HSCT,
557 hematopoietic stem cell transplantation ; SCIG, Subcutaneous immunoglobulin ; UTI, urinary tract infection ; CMV,
558 cytomegalovirus ; y., year(s) ; m., month(s) ; o., old ; + yes or positive, - no or negative.
559



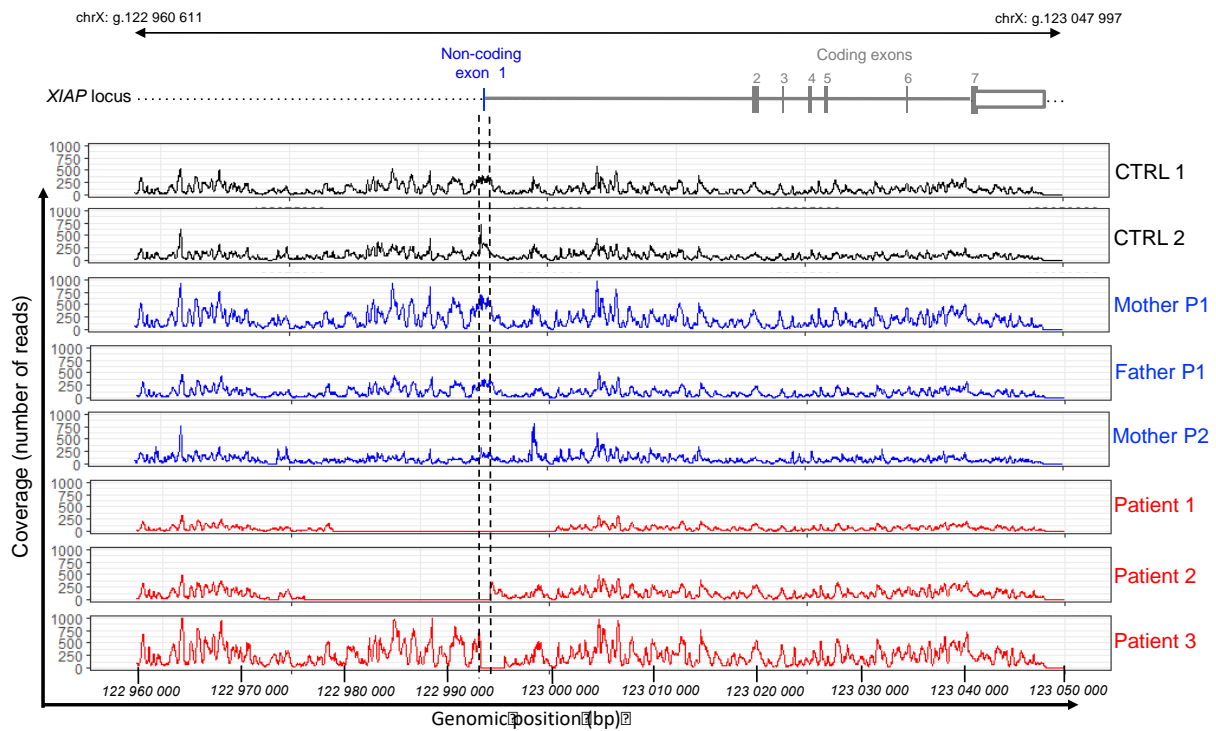
560

561 **Figure 1. Defective XIAP expression in three patients with clinical-like features of XIAP**
 562 **deficiency.**

563 **(A)** Pedigrees of the three unrelated families presenting XLP-2-like features. Black boxes
 564 represent affected individuals. An identification number assigned to each patient is indicated
 565 (P1, P2 and P3).

566 **(B)** Expression of XIAP protein by western blot in T-cell blasts and PBMCs from a healthy
 567 control (CTRL) and P1 (left panels); in the K562 cell line and T-cell blasts from a healthy
 568 control (CTRL) and the P2 (right panels). Blots of KU80 and ACTIN as loading controls. **(C)**

569 Expression of *XIAP* transcripts by PCR in T-cell blasts from CTRL, P1 and P2. GAPDH was
 570 measured in the same samples as a loading control for PCR.



572

573

574 **Figure 2. Identification of large deletions encompassing the non-coding first exon of**575 ***XIAP* in 3 patients with *XIAP* deficiency-like features.**576 Sequencing coverage obtained by NGS of the *XIAP* locus (g.122960611 to g.123047997) of

577 the three patients (Patient 1, Patient 2, and Patient 3), both parents of patient 1 (Father P1 and

578 Mother P1), mother of patient 2 (Mother P2) and two healthy controls (CTRL 1 and CTRL 2).

579 Vertical pale gray lines highlight the position of the exons. Red shading indicates deletions.

580 Vertical dashed lines indicate the common deleted region in the 3 patients encompassing the

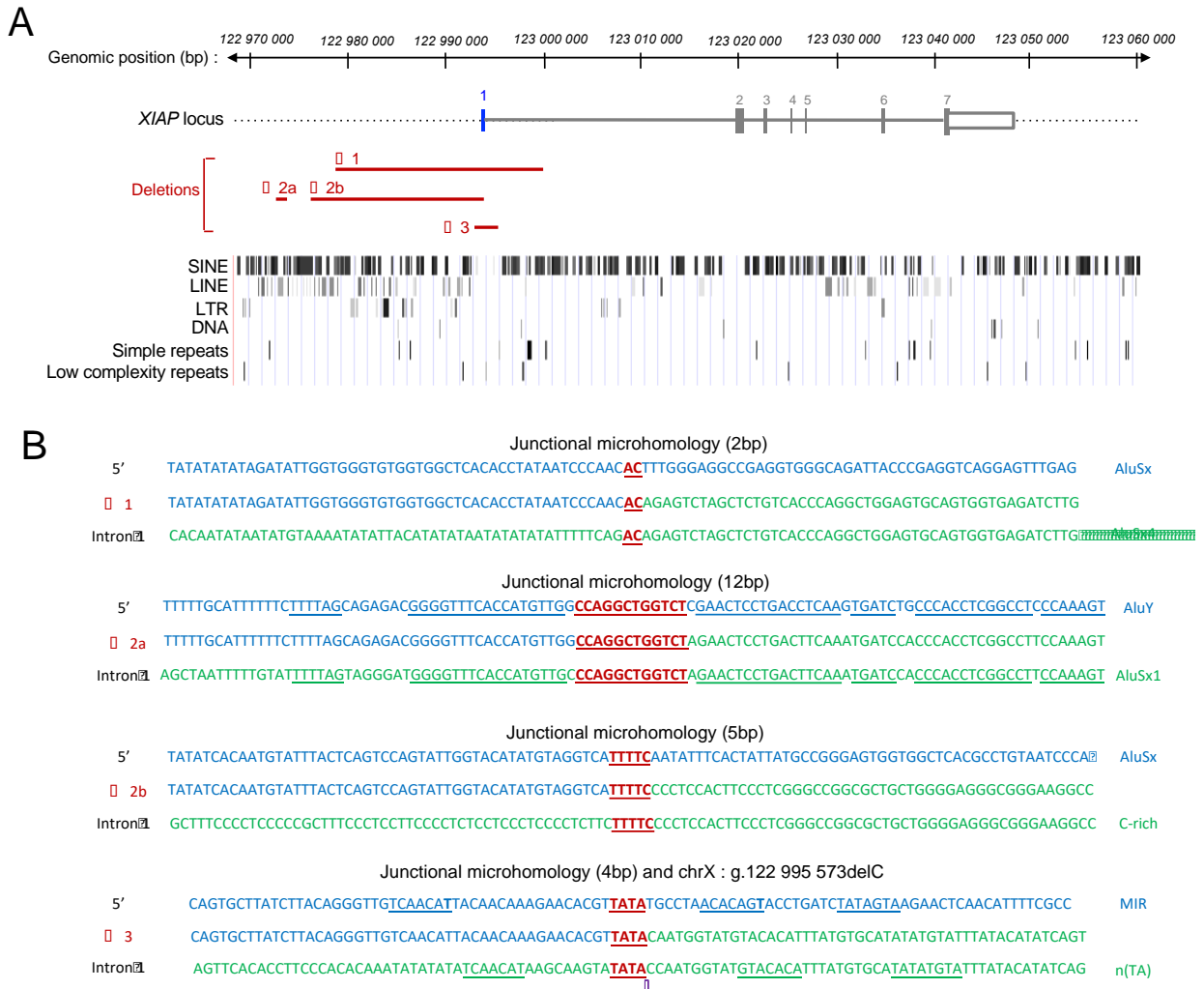
581 non-coding exon1. The *XIAP* locus and the *XIAP* gene structure is depicted on the top with

582 coding exons (grey boxes) and the first non-coding exon (blue box) joined by introns (grey

583 lines). Genomic positions from GRCh37-hg19 human genome assembly.

584

585



587

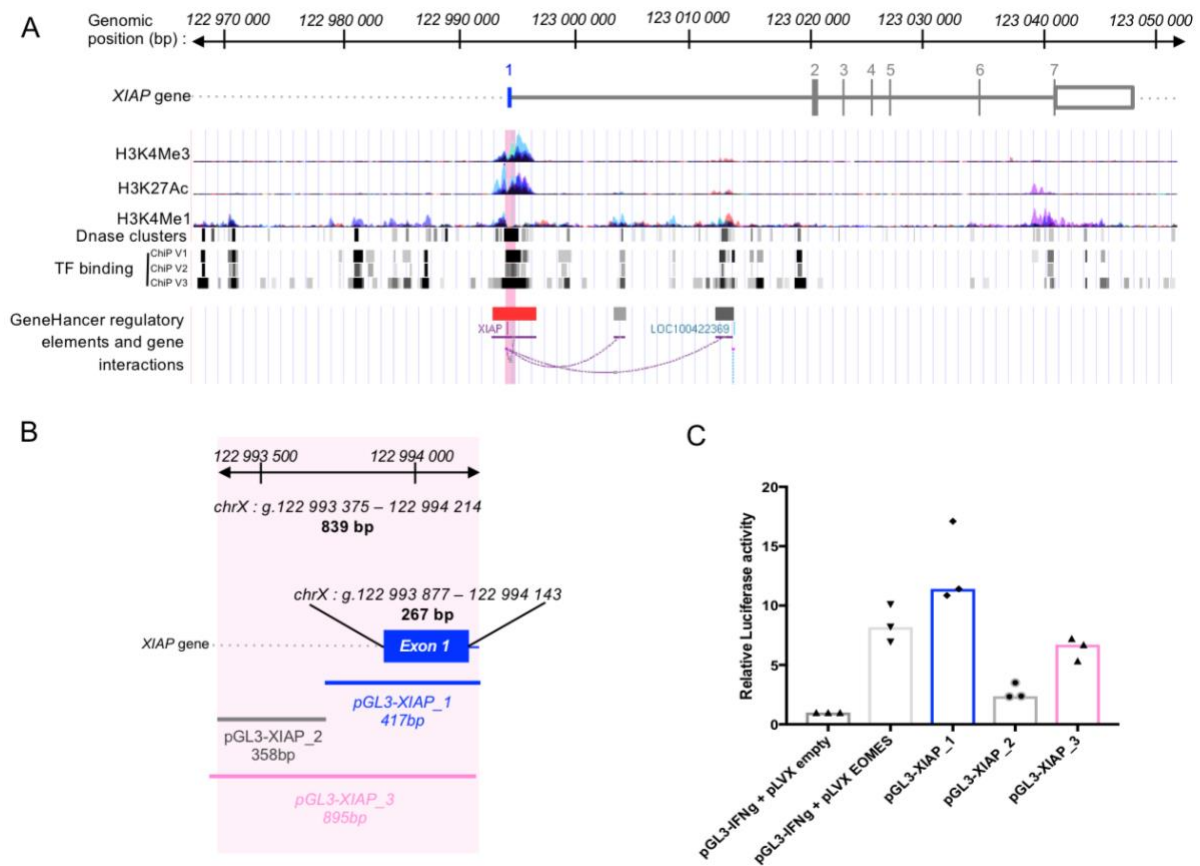
588

589 **Figure 3. Analysis of repeat elements at the *XIAP* locus and mechanisms underlying the three**
 590 **deletions.**

591 (A) Different classes of repeating elements in the *XIAP* locus (g.122968659 to g.123062152) from
 592 ENCODE data bases represented as rectangles/vertical lines, including: short interspersed nuclear
 593 elements (SINE) that include ALUs, long interspersed nuclear elements (LINE), long terminal repeats
 594 (LTR), DNA repeat elements (DNA). Simple repeats include micro-satellites, satellites and RNA
 595 repeats, while low complexity repeats include DNA transposons n(TA). The intensity of the gray
 596 shading in the graphical display inversely reflects the amount of base mismatch, base deletion, and base
 597 insertion associated with a repeated element. The deletions denoted as Δ1, Δ2a, Δ2b and Δ3 identified
 598 in patient 1, 2 and 3 respectively, are represented by red lines. Genomic positions from GRCh37-hg19
 599 human genome assembly.

600 (B) Junctional micro-homologies in red at the breakpoints of the deletions. For each deletion, the 5'
 601 sequences (upper sequences in blue) and sequences in the intron (lower sequences in green) where the
 602 deletion occurred are shown with the deleted sequence corresponding to the middle sequence. Flanking
 603 homologous sequences are underlined. The purple Δ points to the deleted C in the Δ3 deletion.
 604

605



607

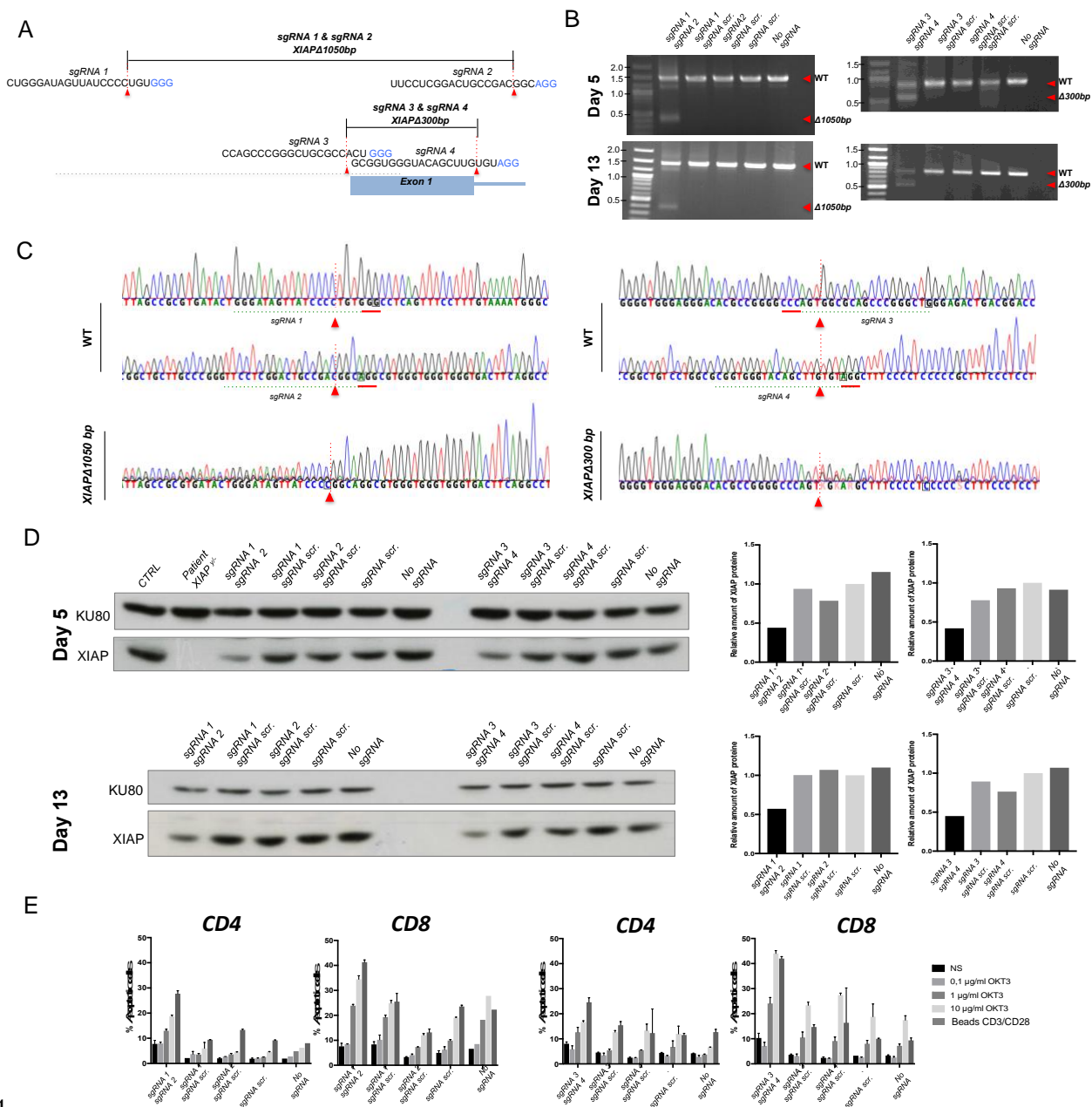
608

609 **Figure 4. The common deleted region in the three patients is an active promoter sequence.**

610 (A) Analysis of the common deleted region in the three patients containing the first non-coding exon
 611 of *XIAP* for regulatory elements and chromatin modifications from UCSC genome browser of the
 612 ENCODE project. Histone modifications H3K4Me3, (indicative of promoters), HEK27Ac (indicative
 613 of active enhancers) and H3K4Me1 (indicative of regulatory regions) from seven different cell lines are
 614 shown in this track. Each cell line is represented by a particular color. DNA clusters corresponding to
 615 DNaseI hypersensitivity (indicative of open chromatin) and binding of transcription factors (TF) from
 616 different ChIP data bases are shown by gray and black boxes/vertical lines that indicates a hypersensitive
 617 region or a peak cluster of transcription factor occupancy. The darkness is proportional to the maximum
 618 signal strength observed in different cell lines tested. GeneHancer Regulatory Elements indicative of
 619 elements with enhancers (in gray) and promoters (in red), and their inferred target genes connected by
 620 curves are indicated at the bottom of the panel. Gray and black boxes correspond to weak and strong
 621 enhancer regions, respectively. The gene structure of *XIAP* is depicted on the top with genome scale of
 622 the *XIAP* locus (as shown in **Figure 2**) from hg19 human genome assembly. The common deleted region
 623 is highlighted by a large pink vertical line.

624 (B) Schematic representation of the commonly deleted region of 839bp sub-divided in two regions of
 625 417bp (containing the non-coding exon1) and 359bp containing the flanking 5' sequences that were
 626 tested in the pGL3-basic (firefly luciferase) vector.

627 (C) Relative dual luciferase activity data from HEK293 cells co-transfected with pRL-CMV (green
 628 *Renilla* luciferase) and pGL3-basic (red firefly luciferase) with a 895bp fragment containing the 839bp
 629 deleted region (pGL3-XIAP_3), the 417bp (pGL3-XIAP_1) or the 358bp pGL3-XIAP_2) sequences as
 630 depicted in (B). Positive control activity corresponds to EOMES-induced activity of the proximal IFNg
 631 promoter. Negative control activity corresponding to the basal activity of the proximal IFNg promoter
 632 (without EOMES) was normalized to 1, and the relative luciferase activity of all the sequences was
 633 expressed with respect to this normalized negative control. Data from three independent experiments.



634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655

Figure 5. Dual-targeting by CRISPR-Cas9 to excise the XIAP non-coding exon1 in primary T cells leads to the decrease of XIAP protein expression. (A) Schematic representation of CRISPR-Cas9-mediated deletions of 105bp ($XIAP\Delta1050bp$) and 300bp ($XIAP\Delta300bp$) encompassing the 839bp region (shaded in pink) and the non-coding exon1 (blue box) respectively. The sequences of sgRNAs (sgRNA1, sgRNA2, sgRNA3 and sgRNA4) are indicated with the PAMs in blue and cleavage sites highlighted with red arrow. (B) PCR analysis over the target sites in *XIAP* at two time points (day 5 and day 13 after transfection) of genomic DNA of primary T cells transfected with the CAS9 and both sgRNA1 and sgRNA2 (sgRNA1/sgRNA2) (left panels), sgRNA3 and sgRNA4 (sgRNA3/sgRNA4) (right panels) or each single sgRNA in the presence of a scramble sgRNA (sgRNA scr.) or without sgRNA (no sgRNA). Wild-type (WT) and $XIAP\Delta1050bp$ amplicons (left panels) for sgRNA1/sgRNA2 are expected at 1453bp and 403bp respectively, while the WT amplicon and $XIAP\Delta300bp$ (right panels) for sgRNA3/sgRNA4 are expected at 843 bp and 543 bp respectively. WT, $XIAP\Delta300bp$ and $XIAP\Delta1050bp$ are depicted by red arrows on the right. (C) Chromatograms from Sanger sequencing of PCR products of panel B showing the wild-type, the $XIAP\Delta1050bp$ and the $XIAP\Delta300bp$ sequences. Breakpoints shown by red arrows. (D) Western blots (on the left) of lysates from primary T cells of panel B. A lysate from XIAP-deficient patient (Patient $XIAP^{P/-}$) was analyzed as a negative control. Blots of KU80 as loading controls. Graphs (on the right) from densitometry quantifications of XIAP expression normalized to KU80 western blots. (E) Activation-induced cell death of primary T cells of panel B, in response to anti-CD3 stimulation with OKT3 at different concentrations or anti-CD3/CD28-coated beads.