Genomic architecture at the Incontinentia Pigmenti locus favors de novo pathological alleles through different mechanisms

Francesca Fusco1, Mariateresa Paciolla1,2, Federico Napolitano1, Alessandra Pescatore1, Irene D’Addario3, Elodie Bal4, Maria Brigida Lioi2, Asma Smahi4, Maria Giuseppina Miano1 and Matilde Valeria Ursini1,*

1Institute of Genetics and Biophysics ‘Adriano Buzzati-Traverso’, IGB-CNR, Naples 80131, Italy, 2University of Basilicata, Potenza 85100, Italy, 3Dermopatic Institute of the ‘Immacolata’, IRCCS, Rome 00167, Italy and 4INSERM U781 et Département de Génétique, Hopital Necker-Enfants Malades, Paris 75015, France

Received August 22, 2011; Revised October 13, 2011; Accepted November 22, 2011

IKBKG/NEMO gene mutations cause an X-linked, dominant neuroectodermal disorder named Incontinentia Pigmenti (IP). Located at Xq28, IKBKG/NEMO has a unique genomic organization, as it is part of a segmental duplication or low copy repeat (LCR1–LCR2, >99% identical) containing the gene and its pseudogene copy (IKBKGP). In the opposite direction and outside LCR1, IKBKG/NEMO partially overlaps G6PD, whose mutations cause a common X-linked human enzymopathy. The two LCRs in the IKBKG/NEMO locus are able to recombine through non-allelic homologous recombination producing either a pathological recurrent exon 4–10 IKBKG/NEMO deletion (IKBKGdel) or benign small copy number variations. We here report that the local high frequency of micro/macro-homologies, tandem repeats and repeat/repetitive sequences make the IKBKG/NEMO locus susceptible to novel pathological IP alterations. Indeed, we describe the first two independent instances of inter-locus gene conversion, occurring between the two LCRs, that copies the IKBKG pseudogene variants into the functional IKBKG/NEMO, causing the de novo occurrence of p.Glu390ArgfsX61 and the IKBKGdel mutations, respectively. Subsequently, by investigating a group of 20 molecularly unsolved IP subjects using a high-density quantitative polymerase chain reaction assay, we have identified seven unique de novo deletions varying from 4.8 to ~115 kb in length. Each deletion removes partially or completely both IKBKG/NEMO and the overlapping G6PD, thereby uncovering the first deletions disrupting the G6PD gene which were found in patients with IP. Interestingly, the 4.8 kb deletion removes the conserved bidirectional promoterB, shared by the two overlapping IKBKG/NEMO and G6PD genes, leaving intact the alternative IKBKG/NEMO unidirectional promoterA. This promoter, although active in the keratinocytes of the basal dermal layer, is down-regulated during late differentiation. Genomic analysis at the breakpoint sites indicated that other mutational forces, such as non-homologous end joining, Alu-Alu-mediated recombination and replication-based events, might enhance the vulnerability of the IP locus to produce de novo pathological IP alleles.

INTRODUCTION

Incontinentia Pigmenti (IP; OMIM 308300) is a rare X-linked dominant disease, lethal in males, affecting the neuroectodermal tissues, always associated with skin defects. The cutaneous lesions, starting in the neonatal period and naturally evolving in four successive inflammatory stages, are hallmarks for IP diagnosis (1,2).

About 73% of IP females carry a loss-of-function mutation in the IKBKG/NEMO gene (inhibitor of kappa light polypeptide...
gene enhancer in B-cells, kinase gamma/nuclear factor kappaB, essential modulator, GenBank NM_003639.3, OMIM 300248), whereas in 27% of cases the mutation still remains elusive (3).

The IKBKG/NEMO gene encodes for NEMO/IKKgamma which acts as a regulatory subunit of the inhibitor of the kappab (IkB) kinase (IKK) complex required for canonical nuclear factor kappaB (NF-kappab) pathway activation involved in many fundamental physiological functions (4,5).

It is currently believed that the highly heterogeneous and often severe clinical presentation of IP might be due to the pleiotropic role of the NEMO/IKKgamma. In addition, the absence of NEMO/IKKgamma protein makes the cells sensitive to apoptosis, leading to the IP-associated male lethality and skewed X-inactivation in females (6–10).

The IKBKG/NEMO locus has a unique genomie organization. In the centromeric direction, the IKBKG/NEMO gene partially overlaps the G6PD (glucose-6-phosphate dehydrogenase, GenBank NM_000402) gene. The two overlapping genes share a conserved promoter region (promoterB), which has a housekeeping bidirectional activity (11–13). It is worth noting that G6PD is a disease gene, causing the X-linked G6PD deficiency (OMIM 305900) (14), the most common enzymopathy in humans. More then 140 different hypomorphic mutations of G6PD have been identified so far, whereas large deletions or a loss-of-function mutation of this gene have never been discovered, leading geneticists to consider that the absence of this gene might be lethal in males (15–17).

In the telomeric direction, the IKBKG/NEMO gene is part of a 35.7 kb segmental duplication containing two low copy repeats (LCRs) arranged in an opposite orientation, one covering the functional gene (LCR1) and the other its partial pseudogene copy (LCR2).

IKBKG/NEMO is structured in nine coding exons (exons 2–10) and four alternative non-coding first exons. Transcription from exons 1B and 1C is directed by the strong bidirectional promoterB, whereas that from exons 1D and 1A is directed by a weak unidirectional promoterA, located in intron 2 of G6PD (13). The non-functional IKBKG spans the region between exons 3 and 10.

A recurrent deletion (IKBKGDel) frequently associated with IP (>70% of cases) is generated by non-allelic homologous recombination (NAHR) due to a misalignment between two Medium Reiterated 67B (MER67B) repeated sequences. The deletion removes a region, ~11.7 kb in length, spanning exons 4–10 in the IKBKG/NEMO gene. Furthermore, the NAHR mechanism can also generate benign copy number variations (CNVs), such as the exon4–10 IKBKG/ deletion, named IKBKGDel, and the exon4–10 IKBKG/NEMO gene duplication, named MER67Bdup, that are both risk alleles for pathological IKBGd in IP patients. Notably, 65% of IP cases are sporadic carrying a de novo IKBG/NEMO mutation (18,19).

The LCR1 and LCR2 sequence homology makes them prone to recombination, as extensively documented in previous reports. Aradhya et al. (18) detected evidence for sequence exchange between the LCR1 and LCR2 copies pointing out that inversion events promoted by their opposite orientation might be responsible for the maintenance of their similarity.

More recently, we have observed that recombination events produced by NAHR between the two LCRs repositioned an exon4–10del from the pseudogene to the IKBG/NEMO gene, thereby causing the IP pathogenic mutations (19).

Finally, recent papers have provided additional proof showing that recombination between these two LCRs might cause the copy number gain observed in mental retardation families (20,21).

We here report that the high frequency of micro/macro-homologies, tandem repeats and repeat/repetitive sequences might predispose the IP locus to novel types of genomic rearrangements arising through different mechanisms generating pathological alterations (22).

We report the first instances of IKBKG/NEMO pathogenic mutations resulting from gene-conversion events and arising between the two homologous LCRs. Furthermore, using a custom-designed high-density quantitative real-time polymerase chain reaction (qRT-PCR) in a group of molecularly unsolved IP patients, we have identified seven different deletions involving not only IKBKGNEMO, but also the overlapping G6PD gene or, in one case, eliminating the bidirectional promoterB. The analysis of such a promoterB deletion has revealed novel unsuspected features of the regulatory region of the IKBG/NEMO gene, which may explain the manifestation of the skin defects in the carrier patient.

RESULTS

Gene conversion mediated by LCRs in the IP locus causes a de novo mutation in the IKBG/NEMO gene

A mechanism such as inter- or intra-chromatid unidirectional gene conversion might be expected to generate pathological alterations in the IP locus because of the high degree of homology between the two LCRs (>99% identity). Supporting this assumption, we identified two unrelated IP families in which a sequence variation in IKBG NEMO in the unaffected father was unidirectionally copied in IKBG/NEMO by a gene-conversion event causing the disease in the affected daughter.

In the IP-603 family, the proband (Fig. 1A, II:1) carried a known c.1167delC mutation in IKBG/NEMO exon 10, leading to the frameshift p.Glu390ArgfsX61 in the NEMO/ IKKgamma protein (3) and the non-pathogenic c.1167delC in its IKBG/NEMO pseudogene (Fig. 1B). The unaffected father (Fig. 1A, I:2) carried the c.1167delC in the IKBG/NEMO pseudogene (Fig. 1B).

Genotype analysis of the family members performed using five polymorphic microsatellites [short tandem repeats (STRs)] surrounding the IP locus (DXS1684, DXS8061, DXS15, DXS1073 and DXS1108) did not identify any recombination event and revealed that the proband inherited the X chromosomes matching the relative haplotype asset present in each parent (Fig. 1A). Therefore, c.1167delC is located in the inherited paternal allele, suggesting that a gene-conversion event had occurred during paternal meiosis.

In the second family, IP-583 (Fig. 2A), we investigated a sporadic IP case (Fig. 2A, III:1) carrying the common pathological IKBGd and also two non-pathogenic CNVs, IKBG/ and MER67Bdup (Figs. 2A and B and 3A). To assess the mechanism generating the occurrence of the IKBGd, we studied the inheritance of both CNVs in IP-583 relatives and we observed the presence of the
IKBKGPdel in the father (II:2, Fig. 2A and B) and the MER67Bdup in the unaffected mother (II:1) and grandmother (I:2, Fig.2A and B). STR analysis along the female lineage of the IP proband (III:1) revealed the presence of MER67Bdup in the recombinant N2 haplotype, indicating that this duplication is located in the IKBKG/NEMO gene of the maternal allele (Fig. 2A). Consequently, the pathological IKBKGdel is located in the paternal haplotype (N5), suggesting that a gene-conversion event is the most obvious mechanism able to explain the presence of the exon 4–10 deletion in both the pseudogene and the gene in the proband. This may have occurred between the IKBKGdel and the wild-type IKBKG/NEMO gene in the father’s germline homogenizing the gene and pseudogene sequences and making both mutated.

Novel deletions in the IP locus involve IKBKG/NEMO and overlapping genes

Mutations of the IKBKG/NEMO gene are responsible for 70–80% of IP. The lack of IKBKG/NEMO mutations in the remaining 20–30% of IP cases is probably due to the detection system. For this reason, we decided to test 20 patients, with an unambiguous clinical diagnosis of IP (based on skin biopsy analysis) but lacking molecular diagnosis, by using a custom-designed high-density quantitative real-time polymerase chain reaction (qRT-PCR), which proved to be effective in revealing quantitative DNA alterations in the IP locus.

We designed 13 specific primer pairs (Ps): 9 able to amplify fragments with a single localization (PF, P1, P2, P3, P4, P9, P10, P11 and P12; Fig. 3B) covering the region from the G6PD to CTAG2 (Cancer/Testis AntiGen 2, NM_020994, OMIM 300396; Fig. 3B) genes, outside the LCRs; and 4 (P5, P6, P7 and P8) with a double localization, in LCR1 and LCR2 (Fig. 3B, Supplementary Material, Fig. S1A) (19). As a control, we established the genomic profile of the PF to P12 probe in 10 control DNA female samples (XX) and 10 male samples (XY). No apparent alteration in the 13 IP samples was observed, whereas an abnormal RT-PCR genomic profile, found in seven samples (IP-48, IP-51, IP-11, IP-47, IP-50, IP-14 and IP-43), indicated the presence of different size deletions (Supplementary Material, Table S1). In IP-48, both LCR1 and LCR2 were deleted. Indeed, the probes localized in the duplicated region were calculated to have two fewer copies. In the other six IP samples, the quantification probes indicated only the presence of LCR2, whereas in IP-43 two intact LCRs were counted (Fig. 3C, Supplementary Material, Table S1). Thus, we identified seven deletions in heterozygotes of different lengths,
each including both IKBKG/NEMO and the overlapping G6PD genes. To map the breakpoints, we designed additional primers in the 5′ (proximal) and 3′ (distal) sequences flanking the deleted regions (Supplementary Material, Tables S2 and S3). Surprisingly, although the distal junctions (3′_IP) of the seven deletions mapped in different positions within the locus, the proximal junctions (5′_IP) were clustered in two regions that we called R1 and R2 (Fig. 3B). In IP-48 and IP-51, the 5′-breakpoint (5′_bkp) was predicted to be in a region located in G6PD intron5 (R1 region, Fig. 3B, Supplementary Material, Table S2), whereas in IP-11 and IP-47, the 5′-bkp was predicted to be in G6PD 3′ untranslated region (UTR; R1 region, Fig. 3B, Supplementary Material, Table S2). In IP-14 and IP-43, the 5′-bkp was proximal to G6PD exon2 (R2 region, Fig. 3B, Supplementary Material, Table S2), whereas in IP-50 it was in G6PD intron2 (R2 region, Fig. 3B, Supplementary Material, Table S2). The 3′-bkp positions were unique and scattered along LCR1 with the exception of IP-48 and IP-43 whose 3′-bkp positions were located outside LCR1 and LCR2. In IP-48, the 3′-bkp mapped in the region between IKBKGp and CTAG2, whereas in IP-43 the distal breakpoint mapped in the region between IKBKG/NEMO intron1C and intron2 (Fig. 3B, Supplementary Material, Table S3).

The genomic architecture flanking the breakpoint junctions reveals a local high content of repeat elements

To explore why the IKBKG/NEMO locus and the surrounding genomic regions are prone to rearrangement, it was essential to determine the positions of the breakpoint-clustering regions and to analyse the junction-sequence signatures in detail. Therefore, we performed an in silico analysis of the IP locus to search for repeat sequences potentially able to lead to genomic instability (23–30).

The tandem repeats finder (31) showed that the highest number of tandem elements was found to be located in LCR1 and LCR2 (chrX:153 790–153 800 kb and chrX:153 860–153 870 kb; UCSC Genome Browser on Human Feb. 2009 GRCh37/hg19; Supplementary Material, Table S4) as well as in the region upstream of the IKBKG/NEMO gene (chrX:153 750–153 760 kb; Supplementary Material, Table S4).

Repeated sequence analysis, performed by Repeated Mask, revealed an enrichment of short interspersed element (SINE) repeat elements (Alu sequences) in the region proximal to LCR1 (chrX:153 740–153 780 kb), of long-interspersed elements (LINEs) in the region near LCR2 (chrX:153 880–153 900 kb) and of long terminal repeats (LTRs) between LCR1 and LCR2 (chrX:153 820–153 840 kb; Supplementary...
Material, Table S4) and near LCR2 (chrX:153 880–153 900 kb; Supplementary Material, Table S4).

Taken together, these observations reveal that the breakpoint ‘hot spots’ clustered in R1 (chrX:153 740–153 770 kb) and R2 (chrX:153 760–153 780 kb) were in regions with a high content of SINEs (33% in R1 and 28.9% in R2) and tandem repeat elements (24 and 7% in R1 and R2, respectively) but with a low content of LINEs (4.2% in R1 and 14.1% in R2) and LTRs (5.8% in R1 and 6.2% in R2) (Supplementary Material, Table S4). On the contrary, an enrichment of specific repetitive sequences was not observed within the genomic region surrounding the 3′ junctions (Supplementary Material, Table S4).

Finally, we looked for sequence homologies (23–30) in a 10 kb region surrounding the 5′ and 3′ breakpoint junctions in each IP sample. We found that the 11 breakpoint regions were located in AluS elements (four AluScs, two AluSq2s and one AluS2), and AluYs and AluJbs (Table 1). More interestingly, their junction regions appeared to contain microhomology sequences (16–118 bp in length) (Table 1).

**Fine mapping of the breakpoint junctions in the IP-43 sample defines the novel IKBKGdelB IP allele**

The determination of the junction sequences, although it represents a key point in the study of the origin of rearrangements, is not always easy to obtain in the case of large alterations (30), especially where a high sequence complexity surrounds the breakpoints making it very difficult to sequence them. This is the case in our study in which the breakpoint deletions are located in regions with a high content of DNA repeats and, consequently, we were not able to identify the precise break junctions except for the rearrangement found in the IP-43 sample. Indeed, we used different nested primer pairs to narrow down the deleted region observed in the IP samples but we obtained the informative amplicon only in IP-43 (Supplementary Material, Fig. S2). In this case, we amplified a fragment of 5.8 kb in the control DNA and a fragment of 1.0 kb in the IP sample (primers B–E; Supplementary
Table 1. List of IP patients and their observed breakpoints in the IP locus

<table>
<thead>
<tr>
<th>ID</th>
<th>Origin</th>
<th>Minimal deletion position (bp)</th>
<th>Upstream X-chromosomal position</th>
<th>Downstream X-chromosomal position</th>
<th>Deleted region</th>
<th>Extended homology</th>
<th>Micro-homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP-48</td>
<td>Italian</td>
<td>11 5198</td>
<td>153 763 394</td>
<td>153 878 592</td>
<td>G6PD exons 1–4, IKBKG/NEMO, CTAG1A, CTAG1B, IKBKGP</td>
<td>AluY (241 bp)/AluY (290): 84%</td>
<td>16 bp (95%)</td>
</tr>
<tr>
<td>IP-51</td>
<td>Greek</td>
<td>65 249</td>
<td>153 763 394</td>
<td>153 828 643</td>
<td>G6PD exons 1–4, IKBKG/NEMO, CTAG1A</td>
<td>AluY (261 bp)/AluY (298 bp): 88%</td>
<td>70 bp (75%)</td>
</tr>
<tr>
<td>IP-11</td>
<td>Italian</td>
<td>63 275</td>
<td>153 757 815</td>
<td>153 821 090</td>
<td>G6PD, IKBKG/NEMO, CTAG1A</td>
<td>AluS (292 bp)/AluSc (315 bp): 75%</td>
<td>57 bp (85%)</td>
</tr>
<tr>
<td>IP-47</td>
<td>Spanish</td>
<td>45 745</td>
<td>153 757 815</td>
<td>153 803 560</td>
<td>G6PD, IKBKG/NEMO</td>
<td>AluSc (226 bp)/AluY (246 bp): 86%</td>
<td>16 bp (95%)</td>
</tr>
<tr>
<td>IP-50</td>
<td>Mexican</td>
<td>26 568</td>
<td>153 769 487</td>
<td>153 796 055</td>
<td>G6PD exons 1 and 2, IKBKG/ NEMO</td>
<td>AluSc5 (382 bp)/AluSc2 (411 bp): 87%</td>
<td>90 bp (85%)</td>
</tr>
<tr>
<td>IP-14</td>
<td>French</td>
<td>22 218</td>
<td>153 774 414</td>
<td>153 796 632</td>
<td>G6PD exons 1 and 2, IKBKG/NEMO</td>
<td>AluSc5 (382 bp)/AluSc2 (441 bp): 87%</td>
<td>48 bp (95%)</td>
</tr>
<tr>
<td>IP-43</td>
<td>Spanish</td>
<td>~4 800b</td>
<td>153 774 414</td>
<td>153 777 269</td>
<td>G6PD exons 1 and 2, IKBKG/ NEMO exon1B-1C</td>
<td>AluSc (258 bp)/AluSc (299 bp): 86%</td>
<td>118 bp (85%)</td>
</tr>
</tbody>
</table>

All IP patients showed a severe IP phenotype (seizures and mental retardation) but IP-43 had only skin IP alterations. All IP patients were extremely skewed for the X-inactivation pattern in peripheral blood lymphocytes (X-inactivation values ≥90:10%) and in IP-43 also in the skin biopsy.

The deletion extension was calculated by measuring the distance (bp) between the two probes with an altered copy number in the qRT-PCR experiments.

In IP-43, the size of the deletion resulting from the PCR amplification is reported.

Material, Fig. S2). This was the smallest fragment obtained in the IP DNA suggesting the presence of a deletion of about 4.8 kb in IP-43 (Supplementary Material, Fig. S2).

Nucleotide sequence analysis of this fragment (B–E, Supplementary Material, Fig. S2) validated the positional mapping of the deletion but did not allow us to sequence across the breakpoint junctions. Indeed, both the proximal (5′-b kp segment; Fig. 4A and Supplementary Material, Fig. S2A) and distal (3′-b kp segment; Fig. 4A and Supplementary Material, Fig. S2A) breakpoints were located in high-density contiguous Alu elements (88.70 and 85.45% of Alu, respectively; Fig. 4B). This architecture suggested that a Alu-Alu-mediated recombination had occurred in the IP-43 progenitor germline. Two homology regions (HR1 and HR2) were present in 5′-b kp (HR1) and 3′-b kp (HR2) with 86% of identity (Supplementary Material, Fig. S3). In addition, the 5′-b kp contained two long tandem repeats [(TTTG)n (TTT)n, TR in Fig. 4B] making it impossible to sequence their junction fragments.

**IKBKGdelB deletion abolishes the promoterB that regulates IKBKG/NEMO expression in keratinocyte differentiation**

The 4.8 kb deletion detected in the IP-43 sample removed the 5′ UTR end portion of the IKBKG/NEMO gene (covering the non-coding 1B and 1C exons) and the 5′ region of G6PD (covering exons 1 and 2) (Fig. 4A). We have named this novel deletion IKBKGdelB, because it eliminates the entire IKBKG/NEMO promoterB, leaving intact the IKBKG/NEMO exon 2 that contains the ATG. The IKBKGdelB allele in the IP-43 patient would be expected to lack the transcription driven by promoterB, leaving only the mRNA transcribed from promoterA. We have previously shown that the IKBKG/NEMO promoters A and B act independently (13). Nevertheless, in the skin of the IP-43 patient, who presented only the characteristic dermatosis (Table 1), we did not know whether promoterA was functioning. Indeed, any further analysis of the in vivo expression of the IKBKGdelB allele was denied by the skewed X-inactivation profile (≥90:10%) observed in the IP-43 DNA skin biopsy. We previously reported (13) that promoterA directs transcription from the alternative exon 1D (liver-specific) and 1A isoforms, whereas promoterB directs the transcription from the 1B and 1C isoforms (Fig. 5A). To obtain a better insight into the contributions of the promoterA- and promoterB-directed transcriptions in the skin, we analysed their activity by RT-PCR in three conditions: (i) in vivo on RNA from different layers of the skin from healthy donors (Fig. 5B) and (ii) in vitro on RNA from differentiated and proliferating HaCat keratinocyte cell lines (Fig. 5C). By using primer pairs, specific for each IKBKG/NEMO isoform, we discovered that although 1C was not expressed, 1A was reduced during keratinocyte differentiation and only the 1B transcript expression was consistently present. Therefore, promoterA was functioning in the lower undifferentiated layers of the skin and was down-regulated during the differentiation, whereas promoterB was expressed in all layers of the skin (Fig. 5B). Comparable results were obtained in the in vitro differentiated HaCat keratinocyte cell line (Fig. 5C). In both systems, we determined the expression of DeltaNp63alpha, as a marker of skin differentiation (Fig. 5B and C). Moreover, the level of expression of G6PD did not change during keratinocyte differentiation (data not shown).

**DISCUSSION**

We have previously reported that the IP locus can undergo NAHR producing either pathological rearrangements (IKBKGdel-) or benign variants (MER67Bdup and IKBKGPDel) (19). In this study, we highlight that other molecular mechanisms, such as gene conversion, Non-Homologous End-Joining (NHEJ), Microhomology-Mediated End Joining (MMEJ) and Fork
Stalling and Template Switching (FoSTeS), may take place at the IP locus. We have discovered novel IP mutations generated by de novo events during parental gametogenesis whose origin could be due to the peculiar genomic architecture of the IKBKG/NEMO region. Most of them are novel rearrangements whose breakpoint site analysis reveals the disruption of the overlapping G6PD gene and of one of the IKBKG/NEMO regulatory elements that is controlled during skin development.

Different mechanisms operate in the IP locus enhance the vulnerability of IKBKG/NEMO to mutations

Gene conversion, the non-reciprocal exchange of genetic information between homologous DNA sequences, could be implicated in the IP disease. On the basis of haplotype analyses, we here report an intrachromatid gene-conversion event that may occur during paternal spermiogenesis. Different mechanisms operate in the IP locus that enhance the vulnerability of IKBKG/NEMO to mutations. Gene conversion, the non-reciprocal exchange of genetic information between homologous DNA sequences, could be implicated in the IP disease. On the basis of haplotype analyses, we here report an intrachromatid gene-conversion event that may occur during paternal spermiogenesis. These findings were obtained by studying two unrelated families with two sporadic IP females carrying de novo IKBKG/NEMO mutations: the mutations identified in the functional IKBKG/NEMO are copied from the IKBKGP pseudogene. In both cases, the father had a pseudogene polymorphism (c.1167delC in IP-603 I:2; IKBKGdel in IP-583 II:2, Figs 1 and 2) that perfectly matched the mutation identified in the gene in the affected daughters.

In general, gene conversion cannot be formally distinguished from double crossover (32) and, therefore, we cannot exclude the possibility that a NAHR event between misaligned chromatids had occurred in the probands’ parental germelines. However, the replacement of the wild-type sequence with a mutated sequence (c.1167delC or IKBKGdel) in the gene acceptor, copying it from a highly homologous sequence in the pseudogene donor, is an extremely likely event when compared with the occurrence of an identical de novo alteration. It is worth noting that, gene-conversion events producing sequence identity between the IKBKG/NEMO gene and its pseudogene have already been reported by Aradhya et al. (18), but instances of gene-conversion mutations generating the pathological IKBKG/NEMO allele have not been reported until now. Finally, in the above reported cases, the gene-conversion event occurred in paternal germlines (Figs 1 and 2). It is indeed well known that during spermiogenesis, programmed double-strand breaks (DSBs) are induced to facilitate the chromatin remodelling that takes place in the elongating spermatids (33,34) and that proteins involved in the homologous recombination repair machinery are present (35).

Besides gene-conversion mutations, we provide evidence of other mutational mechanisms facilitated by the local genomic architecture and able to trigger IP alleles as de novo rearrangements. We identified seven different de novo deletions in heterozygotes which had arisen in a context of high-density repeat sequences and micro/macro-homologies (16–118 bp, Table 1) that are well-known substrates for different mutational forces such as NHEJ, MMEJ and/or FoSTeS (26–28). In silico analysis of the DNA sequence at the IP locus revealed a high prevalence of repeat sequences, such as SINEs, LINEs and LTRs, especially at the proximal breakpoints (Supplementary Material, Table S4) (36). These repetitive elements, together with tandem repetitive elements, are able to generate DSBs or to provide unusual DNA secondary structures including cruciforms, hairpins, triplexes and tetraplexes (37) that may inhibit DNA polymerization increasing the probability of generating rearrangements. Moreover, all the deletion breakpoints are clustered within or in the proximity of Alu repeat clusters (Table 1), frequently involved in recombination events (38–40). In addition, the presence at the breakpoint junctions of microhomologies due to short homologous DNA stretches (16–118 bp, Table 1) makes more likely mechanisms based on coupled homologous and non-homologous recombination, homologous single-strand
invasion and completion by an MMEJ event. On the other hand, all seven rearrangements arise in a genomic context in which the presence of LCRs might drive the DSBs (41–43) or might stimulate genomic rearrangements without being physically involved, as in the FoSTeS model (44).

Large deletions of \textit{IKBKG/NEMO} found in IP patients extend to overlapping genes

Besides the recurrent intragenic \textit{IKBKG/NEMO} exon 4–10 deletion described in most IP cases, \textit{IKBKGdel}, we report here seven novel large deletions affecting \textit{IKBKG/NEMO} and the centromeric and telomeric located genes, \textit{G6PD} and \textit{CTAG1A/B}, respectively. Each deletion had a proximal junction in the \textit{G6PD} gene deleting from 3′ UTR to exon 1 (Fig. 3C and Table 1). The genomic organization of the IP locus in the centromeric direction shows that the \textit{IKBKG/ NEMO} and \textit{G6PD} genes are overlapping and divergently transcribed from opposite strands of the DNA in a ‘head-to-head’ organization conserved during evolution (12,13).

The \textit{G6PD} gene encodes for the initial enzyme of the pentose phosphate pathway (14). Glucose-6-phosphate dehydrogenase deficiency, an X-linked disease, is the most common human enzyme defect. The patients, most frequently males, exhibit non-immune haemolytic anaemia in response to stress or fava bean ingestion. Heterozygous female carriers of a \textit{G6PD}-deficient allele may present clinical manifestations of \textit{G6PD} deficiency, i.e. favism and severe chronic haemolytic anaemia (45), although, in most cases, they have almost normal \textit{G6PD} levels due to a selective advantage of cells expressing the normal allele (17).

The \textit{G6PD} gene exhibits a remarkable polymorphism in human populations (>400 variants) and almost all mutations are point alterations resulting in amino acid substitutions.
(15,17,46), whereas large deletions or loss of function of this gene have never been discovered. Data from the g6pd(−) mouse embryo suggest that alleles with rearrangements deleterious for the G6PD gene are embryonic lethal, whereas heterozygous carrier female mice may survive due to the strong selection against mutant cells early in haematopoiesis (47). The results obtained in mice are partially in agreement with the results obtained in women who are heterozygous for certain severe human G6PD mutations who present a drastic skewed X-inactivation essentially confined to the blood (46). However, males carrying the same mutation in hemizygous may survive although presenting severe haemolytic anaemia. On the other hand, an extensive skewing of X-inactivation has commonly been observed in the blood and fibroblasts of IP patients as indeed a skewed X-inactivation has been considered a diagnostic criterion of IP before undergoing genetic diagnosis (1,48).

Our study is the first to report large deletions involving the G6PD gene in IP patients. The patients here described did not present any clinical manifestations matching the classic G6PD deficiency leading us to hypothesize that a somatic progressive selection against cells expressing the mutated alleles may occur. Therefore, the deleted IKBKG/NEMO allele is able to trigger IP outcomes in the carrier females, whereas the G6PD deleted allele is silent because selection masks the G6PD deficiency outcome in the haematopoietic cells.

Finally, the CTAG1A (OMIM 300657) and its identical copy CTAG1B (OMIM 300156) genes are located in Xq28, LCR1 and LCR2 (18). CTAG1A/B cancer/testis antigens genes, which were deleted in three out of seven IP patients (IP-48, IP-51, IP-11), are expressed in a variety of malignant tumours but solely in the testis among normal adult tissues (49). Therefore, no additional abnormalities of the IP phenotype due to the contribution of the deletion of the CTAG1A/B genes should be observed in those patients.

Functional relevance of the IKBKGdelB deletion for the IKBKG/NEMO expression in the keratinocytes

The discovery of the first IP rearrangement involving one of the key regulatory elements of IKBKG/NEMO expression gives us an entry point to investigate the IKBKG/NEMO gene transcription regulation. We have already established that IKBKG/NEMO has two main conserved regulatory regions: promoterB is a strong bidirectional promoter which directs transcription from exons 1B and 1C, whereas promoterA is a weak unidirectional promoter, located in intron 2 of G6PD that directs transcription from exons 1D and 1A (13). In the IP-43 patient, we mapped a genomic deletion (IKBKGdelB) of the bidirectional promoterB. The deletion abolished the exons IKBKG_1B, IKBKG_1C, G6PD_exon1 and G6PD_exon2 that are promoterB-dependent, whereas the IKBKG/NEMO transcriptional start site, located in exon 2, was not affected.

The absence of promoterB would result in a severe form of the disease in IP-43, and the patient instead presented only the classic IP skin alterations. The higher degree of X-inactivation skewing which we measured in the peripheral blood and in the skin biopsy of the patient precluded the expression analysis in vivo and suggested a strong disadvantage for cells expressing the IKBKGdelB which would disappear preferentially after X-inactivation. The coordinated regulation of two IKBKG/NEMO promoters offers a model to explain the skin lesion pathogenesis in the IP-43 patient. Indeed, the patient’s keratinocytes expressing the IKBKGdelB allele in the basal layer of the skin could contain the NEMO/IKKgamma protein translated only from the 1A isoform. During keratinocyte differentiation, the physiological decline of the 1A isoform and the lack of 1B synergically contribute to deplete those cells of the IKBKG/NEMO transcripts, to inhibit the NF-kappaB signalling and to cause premature keratinocyte death by apoptosis. According to the model of IP skin pathogenesis, the spontaneous necrotic cell death of some NEMO-deficient keratinocytes triggers the expression of proinflammatory mediators by the neighbouring wild-type keratinocytes resulting in the development of the skin lesions (50,51).

Conclusions

The presence of highly homologous regions (LCR1 and LCR2) and the high density of repetitive sequences may have a genome structure-destabilizing effect and predispose the IP locus to generate novel rearrangements by different mechanisms. The result of these rearrangements are non-redundant pathological alleles whose analysis has unravelled the existence of a large deletion of the G6PD gene but also novel unsuspected regulatory mechanisms of IKBKG/NEMO gene transcription regulation that could be more intricate than previously thought. Taken together, our results lead us to propose that IP belongs to the class of pathological condition also known as genomic disorders (52).

MATERIALS AND METHODS

Subjects

All the IP patients analysed in this study met the 1993 revised criteria for the classification of IP (1,2). All patients’ material was gathered, after receiving the informed consent from the participants, under protocols approved by the Declaration of Helsinki. Each potential participant with IP was interviewed and asked to complete an extensive questionnaire. All medical records of IP-affected family members were reviewed to confirm the diagnosis. All affected females underwent skin biopsy obtained from the limbs, usually from the legs. The cutaneous lesions (erythema, vesicles and pustules in the first stage, followed by verrucous and keratotic lesions in the second stage, linear hyperpigmentation in the third stage and pale, hairless, scarring patches in the fourth stage) were observed in all patients. The results of the skin biopsies led unequivocally to a diagnosis of IP.

Blood samples were collected in ethylenediaminetetraacetic acid tubes, and genomic DNA was extracted using the conventional salt precipitation technique. All patients had an apparently normal karyotype. The subjects were initially screened by the following conventional assays (9,19): the IKBKG/ NEMO gene deletion (IKBKGdel) and IKBKGp pseudogene deletion were investigated by a long-range PCR with the EXPAND Long Template PCR system (Roche, Mannheim, Germany) as described in Bardaro et al. (53). For point mutation screening of coding sequences, the electropherograms
were obtained by direct sequencing using Big Dye Terminator Cycle Sequencing Reactions on an ABI 3100 (PE Applied Biosystems, Lennik, Belgium) on a PCR fragment of each exon and were compared with the genomic sequences from GenBank and from control samples.

STR microsatellite analyses and qRT-PCR

Five microsatellite markers distributed along the Xq28 IP region (DXS1684, DXS8061, DXS15, DXS1073 and DXS1107) were genotyped in the families under study. We used in this analysis all the polymorphic STR markers available in the region, as described in public databases (UCSC, Ensembl, etc.) and also additional dinucleotide repeats developed and characterized in the Xq28 region as previously described (54). We identified and delimited the unrepeated regions in chrX:153 747 962–153 959 172 by RepeatMask, and we designed six primer pairs for real-time experiments outside the duplicated region, as described in Fusco et al. (19). All amplification products (70–100 bp) were mapped in the unrepeated regions and were interspersed at 3–20 kb of sequence intervals (PF-P1, 15.4 kb; P1–P2, 5.6 kb; P2–P3, 5.4 kb; P3–P4, 1.4 kb; P9–P10, 7.5 kb; P10–P11, 2.3 kb; P11–P12, 47.6 kb; P12–P13, 7.2 kb). RT-PCR assays were carried out on the Applied Biosystems (Lennik, Belgium) 7900HT system by using the interaction of SYBR Green as a fluorescent reporter (Power SYBR Green PCR Master Mix, Applied Biosystems). The primers, designed using Primer express 3.0 oligo software (Applied Biosystems), are available upon request. All quantifications were normalized to the endogenous gene control Beta-2-microglobulin (B2M gene, GenBank NM_004048.2; OMIM 109700). A total of 25 ng of genomic DNA (in triplicate) from each sample was subjected to RT-PCR experiments using the specific synthetic primers. The DNA copy number was determined as 2E-ddCt for a patient versus a female control, as described in Fusco et al. (19).

Long-range PCR amplification

Oligonucleotide real-time data were used to initially pinpoint approximate breakpoint positions in the genome. We next designed primers for deletions to amplify the rearrangement breakpoint junctions. Different orientations and combinations of primers were also tested for breakpoint analyses. Long-range PCR was conducted with TaKaRa LA Taq polymerase according to the manufacturer’s protocol.

Cell culture and centrifugal elutriation and in vitro differentiation assays

Human keratinocytes were grown to confluence as described (55,56), trypsinized and collected by centrifugation. The cells were fractionated according to cell size by methods described earlier (57).

For in vitro differentiation assays, HaCaT cells were grown in 0.1% foetal calf serum, and CaCl₂ was added to a final concentration of 1.2 mM. The cells were harvested at 0 and 6 days after calcium induction.

Gene expression analysis by qRT-PCR

Human total RNA from an undifferentiated keratinocyte cell line and at different stages of differentiation was used as a source of cDNA. Reverse transcription was carried out with the Superscript III enzyme (Invitrogen, Life Technologies, Carlsbad, CA) according to the manufacturer’s protocol. The IKBKG/NEMO 5′ UTR variants and the overall expression of IKBKG/NEMO were detected using primers described in Fusco et al. (13). qRT-PCR assays were carried out on the Applied Biosystems 7900HT system by using the interaction of SYBR Green as a fluorescent reporter (Power SYBR Green PCR Master Mix, Applied Biosystems). An automatically calculated melting point dissociation curve generated after every assay was examined to ensure the presence of a single PCR species and the lack of primer–dimer formation in each sample. Standard curves were generated from serial 1/5 dilutions of cDNA. The corresponding RT-PCR efficiency was calculated for each couple of primers. The normalized gene expression was calculated as described in Fusco et al. (13).

ONLINE RESOURCES


UCSC Genome Browser: http://genome.ucsc.edu/index.html/org=Human.


RepeatMasker: http://www.repeatmasker.org/.


EMBL Nucleotide Sequence Database: http://www.ebi.ac.uk/embl/.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

The authors thank the patients and their families, the physicians, the International Incontinentia Pigmenti Foundation (http://imgen.bcm.tmc.edu/IPIF/) and Association Incontinentia Pigmenti France (http://www.incontinentiapigmenti.fr) for contributing to this research study. We wish to acknowledge Stefania Filosa for helpful discussions.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by TELETHON (grant GGP08125) to M.V.U., by Association Incontinentia Pigmenti France (http://www.incontinentiapigmenti.fr) fellowship to A.P. and to E.B. and by Programme Hospitalier de Recherches Clinique (PHRC) to A.S.


