Journal of Dermatological Science 94 (2019) 229–235

Contents lists available at ScienceDirect

Journal of Dermatological Science

journal homepage: www.jdsjournal.com

Original artcle

# Identification of genetic alterations in extramammary Paget disease using whole exome analysis

Yukiko Kiniwa[a], Jun Yasuda[b,c,**], Sakae Saito[b], Rumiko Saito[b], Ikuko N. Motoike[b,d], Inaho Danjoh[b], Kengo Kinoshita[d], Nobuo Fuse[b], Masayuki Yamamoto[b,e], Ryuhei Okuyama[a,*]

[a] Department of Dermatology, Shinshu University School of Medicine, Asahi 3-1-1, Matsumoto, 390-8621 Japan
[b] Tohoku Medical Megabank Organization, Tohoku University, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8573, Japan
[c] Miyagi Cancer Center, 47-1 Aza-nodayama, Medeshima-shiote, Natori, Miyagi 981-1293, Japan
[d] Tohoku University Graduate School of Information Science, 6-3-09 Aza-aoba, Aramaki, Aoba-ku, Sendai 980-8579, Japan
[e] Department of Medical Biochemistry, Tohoku University Graduate School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980-8575, Japan

ARTICLE INFO

*Article history:*
Received 22 March 2019
Received in revised form 26 March 2019
Accepted 26 March 2019

*Keywords:*
Extramammary Paget disease
Exome
*TP53*
*ERBB2*
*PIK3CA*
Chromosomal copy number alteration

ABSTRACT

*Background:* Extramammary Paget disease (EMPD) is a rare cutaneous malignant neoplasm, and the genomic alterations underlying its pathogenesis are unknown.
*Objective:* To identify tumor-specific genomic alterations in EMPD.
*Methods:* Exome analysis was performed in specimens from three EMPD patients, and target amplicon sequencing was done for genes frequently mutated in other adenocarcinomas.
*Results:* Exome analysis revealed recurrent somatic mutations in several genes, including*TP53, PIK3CA*, and *ERBB2*. We identified additional candidate exons by searching the COSMIC database for exons that are frequently mutated in other adenocarcinomas. We obtained 19 exons in 12 genes as candidate exons, and performed target amplicon sequencing in samples obtained from EMPD patients. New somatic mutations in the *TP53* gene were identified in six EMPD patients. Single nucleotide polymorphism analysis revealed multiple chromosomal alterations in three EMPD specimens, and two specimens exhibited amplification of chromosome 12p13 and losses of 3p21–24, 7q22 and 13q12–21.
*Conclusion:* Our comprehensive genetic analysis identified novel genomic alterations, and will inform treatment options for EMPD.



## 1. Introduction

Extramammary Paget disease (EMPD) is a rare cutaneous adenocarcinoma that most commonly affects the genital and the perianal area of elderly patients [1]. Cancerous cells are thought to originate from apocrine glands in the epidermis and subsequently invade the dermis. Metastatic EMPD is difficult to control using cytotoxic agents and radiation therapy, and the development of novel systemic therapies would be very beneficial.

EMPD tumors have been found to harbor mutations in *KRAS*, *NRAS*, and *BRAF* genes as well as in *PIK3CA* and *AKT1* genes, suggesting the involvement of constitutive activation of RAS/RAF and PI3K/AKT pathways in pathogenesis [2]. Furthermore, germline mutations in mismatch repair genes including *MLH1*, *MLH3*, *MSH2*, *MSH6*, and *PMS2*, which may lead to microsatellite instability, were also found [3]. Interestingly, rates of recurrence and secondary malignancies were increased in patients with *MLH1*/*MSH2* mutations. In addition, single nucleotide polymorphisms (SNPs) in the DNA repair gene *XRCC1* were associated with a higher risk of developing EMPD [4]. However, the molecular mechanisms underlying EMPD have not been fully elucidated.

The landscape of genetic mutations in many cancers has been elucidated using next generation sequencing, and these data may represent novel therapeutic targets. However, rare cancers such as EMPD have not been as well characterized. Here, we used whole exome sequencing to investigate genetic alterations in EMPD. Our results provide insight into the role of genetic alterations in EMPD pathogenesis.

* Corresponding author at: Department of Dermatology, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto 390-8621, Japan.
** Corresponding author at: Tohoku Medical Megabank Organization, Tohoku University, 2-1, Seiryo-machi, Aoba-ku, Sendai 980-8573, Japan.
*E-mail addresses:* jyasuda@megabank.tohoku.ac.jp (J. Yasuda), rokuyama@shinshu-u.ac.jp (R. Okuyama).

https://doi.org/10.1016/j.jdermsci.2019.03.006

## 2. Materials and methods

### 2.1. Patient samples

This study was approved by ethics committees at Shinshu University School of Medicine (approved numbers: 2011-363, 2016-570) and Tohoku University School of Medicine (2017-2), and all samples were obtained with informed consent. Specimens with a large proportion of tumor cells were obtained from three EMPD patients (P1, P2, and P3) for whole exome sequencing and SNP array analysis. An additional 15 specimens obtained from 13 EMPD patients were used for selected target amplicon sequencing. All patients had EMPD in the genital area and were treated at Shinshu University Hospital between 2011 and 2016. The clinical information of all patients is shown in Table 1.

### 2.2. DNA preparation

For whole exome sequencing, surgically resected tumor specimens and corresponding blood samples were stored in liquid nitrogen. Genomic DNA was extracted from tumor tissue using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) and from blood using the DNeasy Blood and Tissue Kit (Qiagen). For target amplicon sequencing, surgically resected EMPD tumor specimens and corresponding non-cancerous specimens from normal skin regions in the surgical margin were first fixed in 10% formalin for 24–48 h and embedded in paraffin. Sections were then stained with hematoxylin and eosin to confirm the area of EMPD lesions, and 10 μm thick sections of cancerous and non-cancerous cells were dissected from the formalin-fixed paraffin-embedded (FFPE) tissues. Genomic DNA was extracted using the GeneRead DNA FFPE Kit (Qiagen). The DNA concentration was measured using a biophotometer, and DNA quality was evaluated using agarose gel electrophoresis.

### 2.3. Exome capture and sequencing

Genomic DNA was sheared into fragments of approximately 300 bp by sonication using a Covaris 2000 ultrasonicator (Covaris, Woburn, MA, USA). To capture exon fragments, we used the SureSelect Human All Exon V5 + Regulatory kit (Agilent, Santa Clara, CA, USA) according to the manufacturer's instructions. Exome sequencing was performed with an ion proton sequencer (Thermo Fisher Scientific, Waltham, MA, USA) as previously described [5]. Briefly, we used the GRCh37 (hg19) genome assembly as the reference sequence and alignment was carried out using Torrent Mapping Alignment Program (TMAP) software version 3.6.39. Torrent Variant Caller (TVC) version 3.6.39 was also used for variant calls, and the parameters for variant discovery were set to use as many reads as possible.

### 2.4. Variant calls and filtration for identification of tumor-specific genetic alterations

After obtaining the raw variant calls, extensive filtering of the Variant Call Format (VCF) files was performed to remove potential false positives. In brief, the low-quality variant calls were filtered, the non-targeting regions were removed from the dataset, and VCF files were further filtered using two panels of non-cancerous samples (Panels of Normals; PoNs). One PoN consisted of the merged variant calls of the peripheral blood DNA from three EMPD patients (P1, P2, and P3), and the second PoN consisted of the 3.5KJPN, 3552 normal Japanese whole genome sequencing panel [6]. First, low-quality variant calls (QUAL value < 80) were removed by the use of VCFtools from blood exome variant call data. Second, the three non-cancerous, filtered variant data were merged and a VCF file was generated as a first PoN. Third, the subtracted (tumor exome-specific) variants were annotated using the ANNOVAR software package and a custom database file was added that gives the minor allele frequencies of 3.5KJPN SNPs [6]. In addition, tumor exome data were filtered with GATK variant filtration with the condition of "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0", and the filtered VCF files were subjected to subtraction with the two PoNs. Finally, the tumor-specific variants that were not found in either blood DNA of the three EMPD cases or 3.5KJPN were selected. Furthermore, variants located in so-called "blacklist" genes were removed from the results [7]. The Ts/Tv ratio was calculated with VCFtools.

### 2.5. Selection of target amplicon sequencing

The databases of ICGC (International Cancer Genome Consortium) and COSMIC (Catalogue of Somatic Mutations in Cancer) were analyzed to identify candidate exons with mutational hotspots that are frequently found in adenocarcinomas such as breast, colon, pancreatic, and thyroid cancers. Primer sequences and their chromosomal positions are shown in Supplementary Table 1. The primers were designed by Ampliseq web service

**Table 1**
Clinical information of patients with extramammary Paget disease.

| Sample ID | Age, sex | Tumor lesions | Non-cancerous tissue | Lymph node metastasis | Distal metastasis | Dataset |
|---|---|---|---|---|---|---|
| P1 | 63, male | Primary (mass) | Blood | – | – | Exome, amplicon, SNP array |
| P2 | 79, female | Primary (mass) | Blood | – | – | Exome, amplicon, SNP array |
| P3 | 62, male | Primary (mass) | Blood | + | + | Exome, amplicon, SNP array |
| P3L | 62, male | Metastatic LN | NA | + | + | Amplicon |
| P4 | 71, male | Primary (mass) | Skin | + | + | Amplicon |
| P4L | 71, male | Metastatic LN | NA | + | + | Amplicon |
| P5L | 58, female | Metastatic LN | NA | + | – | Amplicon |
| P6 | 62, male | Primary (mass) | Blood | + | + | Amplicon |
| P7 | 71, male | Primary (mass) | NA | + | – | Amplicon |
| P8 | 87, female | Primary (mass) | NA | – | – | Amplicon |
| P9 | 63, female | Primary (early invasion) | NA | – | – | Amplicon |
| P10 | 89, male | Primary (in situ) | NA | – | – | Amplicon |
| P11 | 87, male | Primary (in situ) | Skin | – | – | Amplicon |
| P12 | 61, female | Primary (in situ) | Skin | – | – | Amplicon |
| P13 | 83, male | Primary (in situ) | Skin | – | – | Amplicon |
| P14 | 72, male | Primary (in situ) | Skin | – | – | Amplicon |
| P15 | 87, female | Primary (in situ) | NA | – | – | Amplicon |
| P16 | 80, female | Primary (in situ) | NA | – | – | Amplicon |

Abbreviations: LN, lymph node; NA, not applicable.

(https://ampliseq.com), and PCR and MiSeq sequencing were performed as previously described [8]. The FASTQ files were analyzed using GATK Mutect2 ver. 4.1.1.0. The PoN was constructed with non-cancerous tissue samples from six patients (P4, P6, P11, P12, P13, and P14) with default options in the GATK Mutect2 package. The following parameters were changed from the default option for variant calls: disable-read-filter was set to MateOnSameContigOrNoMappedMateReadFilter, and tumor-lod-to-emit was set to 20. The germline data were derived from 3.5KJPN [6].

### 2.6. SNP array analysis

Procedures were described in detail previously [5]. The Human Omni 2.5–8 v1.1 DNA Analysis Kit (Illumina, San Diego, CA, USA) was used to obtain SNP data. The default set cluster file was HumanOmni2-5 M-8b1-1_B.egt (Illumina), and a Gen Call Threshold of 0.15 was used for SNP calling. For the SNP call rate, samples with overall call rates over 99% and LogRdev values below 0.2 were used for further analysis.

## 3. Results

### 3.1. Histopathology and whole exome sequencing of EMPD tumors

We used specimens from three patients with EMPD (P1, P2, and P3) for whole exome sequencing. All three cases had a higher rate of tumor cells than typical EMPD specimens; P1 and P3 showed nodular formations and P2 showed a disseminating pattern with complete replacement of tumor cells in the epidermis (Supplementary Fig. 1).

Tumor-specific mutations identified by exome sequencing are summarized in Table 2. The total length of target regions was 114 MB, and the average number of somatic point mutations per megabase was 42.1, 41.0, and 65.5 for P1, P2, and P3, respectively. Although EMPD is considered a microsatellite-stable cancer, we found large numbers of somatic mutations. Similarly, the

**Table 2**
Summary of tumor-specific variants in three patients with extramammary Paget disease.

| Type | P1 | P2 | P3 |
|---|---|---|---|
| Total tumor-specific alterations | 5882 | 5678 | 8677 |
| Exonic | 1332 | 1018 | 1906 |
| Splicing | 13 | 4 | 9 |
| Non-synonymous | 604 | 475 | 945 |
| Stop gain | 31 | 11 | 41 |
| Deletions | 443 | 392 | 407 |
| Insertions | 642 | 612 | 800 |
| Frameshift | 122 | 86 | 117 |
| COSMIC70 | 32 | 35 | 35 |
| COSMIC70 frameshift | 12 | 12 | 16 |
| Average SNA number/megabase | 42.1 | 41.0 | 65.5 |

Abbreviation: SNA, single nucleotide alteration.

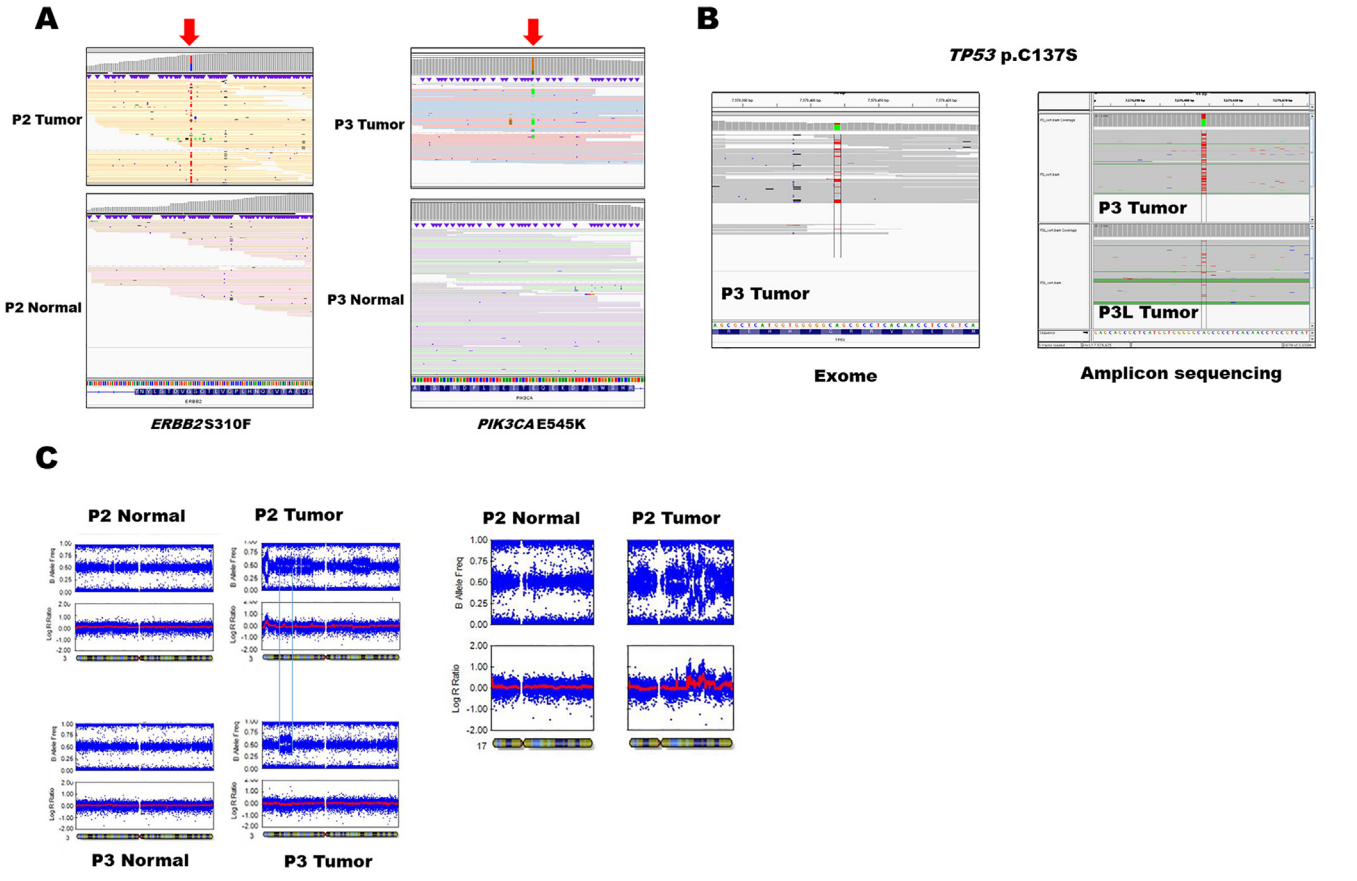


**Fig. 1.** (**A**) *ERBB2* and *PIK3CA* mutations in tumor cells of extramammary Paget disease. The integrative genome viewer (IGV) images of the exome analysis are shown. Panels indicate the corresponding pairs of P2 at *ERBB2* S310 F and P3 at *PIK3CA* E545 K. Red arrows indicate the location of mutations. (**B**) Compensation of mutation calls by amplicon sequencing. The left panel is an IGV image of the exome data at C137S in the *TP53* gene of P3. The right panel shows the corresponding IGV image of P3 obtained by amplicon sequencing. (**C**) Chromosomal copy number alterations in extramammary Paget disease. Left panels: Copy number alterations in chromosome 3 of P2 and P3. Chromosome 3p shows deletions in tumor samples. B allele frequencies and log R ratio are indicated in the upper and lower parts of each panel, respectively. Right panels: Complex copy number alterations in chromosome 17 of P2.

frequency of indels was high (14–18% of the detected tumor-specific alterations) in all three patients. The ion proton sequencer may be error-prone for indel calls, particularly at homopolymer sequences [9–11]. Compared with the genotyping results of SNP arrays, the sensitivity of ion proton exome analysis was 0.828–0.882 and the positive prediction value was 0.796–0.837 (Supplementary Table 2), suggesting that variant calls in ion proton sequencing may include false positives or false negatives. We hypothesize that many of the variant calls actually exist in tumor cells because we were able to verify some variant calls using amplicon sequencing. The Ts/Tv ratio obtained was similar to those from other analyses of the cancer exome, indicating that the variant calls may not be distorted by systemic biases (Supplementary Table 3).

### 3.2. TP53, ERBB2, *and* PIK3CA *mutations are involved in EMPD*

We next compared our ion proton exome dataset with comprehensive gene databases. We searched the COSMIC database for genes, and found 85 genes. Many of these had alterations that were the same as those in our exome sequencing data: 32 in P1, 35 in P2, and 35 in P3 (Table 2).

We identified 18 somatic mutations in genes frequently mutated in other types of cancer (defined as more than five cases in the COSMIC database), including seven pathogenic and two frameshift mutations (Table 3 and Supplementary Table 4). We identified somatic mutations in pan-cancer driver genes such as *TP53*, *PIK3CA*, and *ERBB2*: *TP53* (E286 K) and *ERBB2* (S310 F) were found in P2, and *PIK3CA* (E545 K) was found in P3 (Fig. 1A). In addition, a nonsynonymous alteration (D643E) in *CDK11B* was found in two patients (P1 and P3). However, because the *CDK11* gene is duplicated in the human genome (*CDK11 A* and *CDK11B*), the *CDK11B* mutations detected in this study might be artifacts caused by misalignment of the reads derived from *CDK11 A* (data not shown).

### 3.3. Target amplicon sequencing supports the results of exome analysis

Following cleanup of the exome data, we searched for candidate driver mutations found in other adenocarcinomas. We performed target amplicon sequencing in 18 samples from EMPD patients and six samples from non-cancerous controls. The target genes were chosen from the COSMIC database based on their frequency of mutation in adenocarcinomas. We selected 19 exons in 12 genes; target somatic alterations are shown in Table 4 and Supplementary Tables 5 and 6. The amplicon sequencing results supported our ion proton exome sequencing data, and we found somatic alterations in the *TP53* gene in seven patients (Fig. 1B and Table 4). All alterations specific to cancerous samples are shown in Table 4.

### 3.4. Copy number alterations in EMPD

We investigated chromosomal alterations in three cases of EMPD (P1, P2, and P3) using SNP array analysis. Chromosomal regions with altered copy number differed in the three cases examined (Supplementary Table 7 and Fig. 1C). The alterations were localized in P1, but were distributed in multiple chromosomal regions in P2 and P3. Interestingly, some chromosomal alterations were observed in multiple cases: gain of chromosome 12p13 in P2 and P3; loss of chromosome 7q22 in P1 and P3; and loss of chromosomes 3p21–24 and 13q12–21 in P2 and P3. The presence of these copy number alterations in multiple patients suggests that they are related to pathogenesis of the disease.

Somatic copy number loss is occasionally related to inactivation of tumor suppressor genes. When we compared our results with the list of cancer-predisposing genes used in The Cancer Genome

**Table 3**
Mutations identified in EMPD specimens as well as in more than five cases in the COSMIC70 database.

| Sample ID | Gene | COSMIC annotation | Amino acid changes | FATHMM prediction |
|---|---|---|---|---|
| P1, P3 | *CDK11B* | ID = COSM145214,COSM3360408;OCCURENCE = 1(central_nervous_system), 1(esophagus), 1(thyroid), 1(lung), 2(hematopoietic_and_lymphoid_tissue), 1(large_intestine), 1(kidney) | D385E | Pathogenic |
| P3 | *CR1* | ID = COSM301989,COSM301990;OCCURENCE = 1(large_intestine), 1(prostate), 1(kidney), 3(endometrium), 2(lung), 1(central_nervous_system) | R1744X | Neutral |
| P3 | *DCHS1* | ID = COSM243971;OCCURENCE = 1(endometrium), 5(breast), 1(prostate) | L33_G34insL | NA |
| P3 | *DENND4B* | ID = COSM1319934,COSM1319935;OCCURENCE = 1(ovary), 6(breast) | Q908_Q910delQQQ | NA |
| P2, P3 | *ENTPD2* | ID=COSM3763808;OCCURRENCE=5(thyroid), 1(large_intestine), 1(hematopoietic_and_lymphoid_tissue) | I495I | Neutral |
| P2 | *ERBB2* | ID = COSM48358;OCCURENCE = 9(urinary_tract), 3(large_intestine), 3(breast), 1(ovary), 3(lung), 1(stomach) | S310F | Pathogenic |
| P2 | *ERICH1* | ID = COSN149612;OCCURENCE = 1(liver), 1(kidney), 1(stomach), 2(thyroid) | Intergenic non-coding | NA |
| P2 | *ESRRA* | ID = COSM1746433;OCCURENCE = 4(urinary_tract) | R376L | Pathogenic |
| P3 | *MAGI1* | ID = COSM1048140,COSM1485456,COSM1048138;OCCURENCE = 6(breast), 1(urinary_tract), 1(endometrium) | Q421_T422insQ | NA |
| P1 | *NUP93* | ID = COSM124893;OCCURENCE = 2(urinary_tract), 5(breast), 1(lung), 1(upper_aerodigestive_tract) | E14 K | Pathogenic |
| P3 | *PIK3CA* | ID = COSM763,COSM125370;OCCURENCE = 4(salivary_gland), 38(endometrium), 4(soft_tissue), 28(ovary), 1(pituitary), 4(kidney), 3(pancreas), 3(liver), 126(urinary_tract), 1(biliary_tract), 373(large_intestine), 1(prostate), 365(breast), 16(skin), 1(small_intestine), 4(upper_aerodigestive_tract), 8(esophagus), 18(stomach), 17(central_nervous_system), 21(cervix), 7(gastrointestinal_tract_ (site_indeterminate)), 4(penis), 44(lung), 7(thyroid) | E545 K | Pathogenic |
| P3 | *SKIDA1* | ID = COSM305499,COSM305498;OCCURENCE = 1(hematopoietic_and_lymphoid_tissue), 1(endometrium), 1(ovary), 1(soft_tissue), 1(liver), 5(breast) | E428_G429insEE | NA |
| P2 | *TMEM184A* | ID = COSM1088016;OCCURENCE = 1(endometrium), 2(large_intestine), 4(breast) | G392_S393insG | NA |
| P2 | *TP53* | ID = COSM99924,COSM3522693,COSM10726,COSM1645467;OCCURENCE = 8(upper_aerodigestive_tract), 12(esophagus), 4(central_nervous_system), 5(hematopoietic_and_lymphoid_tissue), 7(lung), 2(ovary), 1(soft_tissue), 5(urinary_tract), 2(liver), 1(biliary_tract), 4(large_intestine), 8(breast), 8(skin) | E286K | Pathogenic |
| P1 | *UHRF1* | ID = COSM1394330;OCCURENCE = 12(large_intestine) | P674fs X35 | NA |
| P3 | *YWHAEP7* | ID = COSM472658;OCCURENCE = 5(endometrium), 1(kidney), 1(NS), 2(cervix) | T41N | Neutral |
| P1 | *ZFHX4* | ID = COSM218965,COSM3395414;OCCURENCE = 3(pancreas), 1(endometrium), 1(lung) | R1983H | Pathogenic |
| P1 | *ZFYVE19* | ID = COSM1323224;OCCURENCE = 4(large_intestine), 2(urinary_tract), 2(ovary) | R48fs X63 | NA |

Abbreviation: NA, not applicable.

**Table 4**
Variants detected by amplicon sequencing.

| Chromosome | 3 | 3 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Position | 178936091 | 178952085 | 7577022 | 7577082 | 7577099 | 7577120 | 7577539 | 7578404 | 37868208 | 2226730 | 60895692 |
| Ref | G | A | G | C | C | C | G | A | C | C | C |
| Alt | A | G | A | T | T | T | A | T | T | T | T |
| Gene.refGene | PIK3CA | PIK3CA | TP53 | TP53 | TP53 | TP53 | TP53 | TP53 | ERBB2 | DOT1L | LAMA5 |
| ExonicFunc.refGene | Missense | Missense | Stopgain | Missense | Missense | Missense | Missense | Missense | Missense | Missense | Missense |
| AAChange.refGene | E545K | H1047R | R306X | E286K | R280K | R273H | R248W | C176S | S310F | P1404S | E2228K |
| COSMIC70 | YES | YES | YES | YES | YES | YES | YES | YES | YES | NO | NO |
| P1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| P2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| P3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| P3L | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P4 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P4L | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| P6 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| P7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P12 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| P13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| P15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| P16 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total positives | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Concordance with exome | YES | NA | NA | YES | NA | NA | NA | NO | YES | NA | NO |

Abbreviation: NA, not applicable.

Atlas (TCGA) project [12], we found that several interesting genes were located in the regions of copy number alteration. However, we did not identify homozygous mutations in the corresponding genes in our whole exome analysis.

## 4. Discussion

This study is the first genome-wide survey of somatic alterations in EMPD. Exome sequencing identified somatic mutations in 9 genes, including seven pathogenic (*CDK11B*, *ERBB2*, *ESRRA*, *NUP93*, *PIK3CA*, *TP53*, and *ZFHX4*) and two frameshift mutations (*UHRF1* and *ZFYVE19*). Amplicon sequencing uncovered somatic mutations in the *TP53* and *PIK3CA* genes in multiple cases of EMPD. Additionally, various chromosomal alterations were identified by SNP array.

In the current study, we identified mutations in *PIK3CA* E545 K and H1047R. The PI3K signaling pathway participates in a broad range of cellular regulatory processes, and aberrations in PI3K signaling contribute to an equally broad spectrum of human diseases, including cancer. *PIK3CA*, which encodes PI3K p110α, is frequently mutated in cancers, and E545 K is a common mutation that disrupts inhibitory interactions with PI3K regulatory molecules [13]. H1047R is another common mutation, and it occurs in breast cancer with especially high frequency. In a mouse model of breast cancer, *PIK3CA* H1047R was reported to support multipotency of mammary tumors [14]. It is likely that the *PIK3CA* mutations we observed, E545 K and H1047R, are significant in EMPD pathogenesis. Although early clinical trials of PI3K pathway inhibitors were not as successful as had been hoped, the PI3K pathway remains a promising target for drug development. For example, taselisib selectively inhibits PI3K p110α, and clinical studies of its use as a treatment for various cancers are ongoing [15]. Development of PI3K pathway inhibitors may provide clinical benefit in some patients with EMPD.

We found a mutation, S310 F, in the receptor tyrosine kinase *ERBB2* (also known as HER2), similar to previous reports [16–18]. Although *ERBB2* mutations in the COSMIC database are chiefly located in the kinase domains, the S310 F mutation is located in the extracellular domain. Because *ERBB2* mutations in the extracellular domain are frequently observed in micropapillary urothelial carcinoma [19], the S310 F mutation is likely oncogenic [20]. Tumors with this mutation may be sensitive to HER2 (ERBB2) inhibitors. A case report has been published describing the clinical effect of lapatinib, a dual tyrosine kinase inhibitor targeting HER2 and EGFR, in a patient with the S310 F mutated EMPD [18].

*TP53* was the most frequently affected gene in the present study. *TP53* mutations had been previously reported to be associated with EMPD [21], and we found *TP53* mutations in 6 of the 18 cases examined (Tables 3 and 4). Here, the R248 W mutation was identified; codon 248 is the most common mutational hotspot in skin cancers, including melanoma and squamous cell carcinoma [22]. However, other mutations that we found (R306X, R280K, R273H, and C176S) have not been found in other skin cancers. Unlike other skin cancers, EMPD occurs in areas not typically exposed to sunlight, suggesting that ultraviolet irradiation does not play a role in EMPD. The location of EMPD, in the genital and perianal areas, suggests that persistent irritation due to skin friction may play a role in development of the disease.

ERRα (estrogen-related receptor alpha), encoded by *ERSSA*, is a member of the nuclear receptor superfamily. High expression of ERRα is associated with a poor prognosis in cancers originating from the colon, ovary, prostate, and breast [23–25]. Interestingly, ERRα plays a role in energy regulation due to mitochondrial oxidative metabolism [26]. The R376 L mutation is located in the hormone-binding region, and is found in urinary tract cancer.

ZFHX4, a transcription factor, was reported to control the glioblastoma tumor-initiating cell state via modulation of epigenetic regulation [27]. In addition, mutations in *ZFHX4* are associated with poor prognosis in esophageal squamous cell carcinoma [28], and the R1983H mutation is observed in pancreatic and other cancers. NUP93 is a subunit of the nuclear pore complex, and regulates molecular transport between the nucleus and cytoplasm. The E14 K mutation is associated with breast and other cancers in the COSMIC database. *UHRF1*, a potent

oncogene overexpressed in various tumors [29,30], affects expression of tumor suppressor genes via epigenetic modification. Further investigation is needed to clarify the contribution of these genes to EMPD.

Exome analysis of cancer specimens is expensive and difficult to perform on a large scale. As a more targeted approach, there have been previous attempts to establish cancer-specific gene panels for identification of somatic mutations [31,32]. We complemented our exome analysis by selecting a panel of cancer-specific genes to examine in additional EMPD cases. Somatic mutations are usually affected by the origin of cancers [12,33], but EMPD is a rare disease that has not been well characterized, so we selected adenocarcinoma-specific target amplicons for our gene panel. Further technical improvements are necessary to utilize this approach in a clinical setting.

Chromosomal copy number alterations have previously been observed in EMPD. Whereas a prior report described amplifications in chromosome 19 and the long arm of the X chromosome together with loss of 10q24–qter [34], we detected new chromosomal alterations in multiple regions (Supplementary Table 7). Interestingly, we observed a similar pattern of chromosomal gains and losses in two cases; tumors from P2 and P3 both had gains of chromosome 12p13 and losses of chromosomes 3p21–24 and 13q12–21. Several cancer-predisposing genes are located in these regions. *ETV6* on chromosome 12p13 is known to join other genes through rearrangement in hematological malignancies [35]. The *BRCA2* and *RB1* tumor suppressor genes are located on chromosome 13q12–21, and loss of this region may contribute to EMPD carcinogenesis. The *MLH1* mismatch repair gene is located on chromosome 3p22.2, and its loss may lead to microsatellite instability [3]. Multiple cancer-predisposing genes are theoretically altered in the three EMPD cases we used for whole exome analysis (Supplementary Table 7); however, we failed to detect deleterious mutations in the exome data. In the case of tumor suppressor genes, structural changes (such as loss of distant enhancer regions) or epigenetic changes may occur in the remaining wild-type alleles.

We performed a comprehensive genetic analysis of EMPD and identified candidates for driver mutations. Additional analysis will clarify the mutational signature of EMPD and lead to improved clinical outcomes.

## Funding sources

This study was supported by JSPS KAKENHI Grant Numbers 16K10149 (Y.K.) and 23501264 (J.Y.).

## Acknowledgments

We thank Ms. Aya Uchiyama, Natsumi Konno, Noriko Takahashi, and Chizuru Abe for technical assistance.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jdermsci.2019.03.006.

## References

[1] J.M. Cohen, S.R. Granter, A.E. Werchniak, Risk stratification in extramammary Paget disease, Clin. Exp. Dermatol. 40 (2015) 473–478.
[2] Z. Kang, F. Xu, Q.A. Zhang, Z. Wu, X. Zhang, J. Xu, Y. Luo, M. Guan, Oncogenic mutations in extramammary Paget's disease and their clinical relevance, Int. J. Cancer 132 (2013) 824–831.
[3] Z. Kang, F. Xu, Y. Zhu, P. Fu, Q.A. Zhang, T. Hu, X. Li, Q. Zhang, Z. Wu, X. Zhang, H. Wang, J. Xu, Z. Fang, M. Guan, Genetic analysis of mismatch repair genes alterations in extramammary paget disease, Am. J. Surg. Pathol. 40 (2016) 1517–1525.
[4] K. Chiyomaru, T. Nagano, C. Nishigori, XRCC1 Arg194Trp polymorphism, risk of nonmelanoma skin cancer and extramammary Paget's disease in a Japanese population, Arch. Dermatol. Res. 304 (2012) 363–370.
[5] I.N. Motoike, M. Matsumoto, I. Danjoh, F. Katsuoka, K. Kojima, N. Nariai, Y. Sato, Y. Yamaguchi-Kabata, S. Ito, H. Kudo, I. Nishijima, S. Nishikawa, X. Pan, R. Saito, S. Saito, T. Saito, M. Shirota, K. Tsuda, J. Yokozawa, K. Igarashi, N. Minegishi, O. Tanabe, N. Fuse, M. Nagasaki, K. Kinoshita, J. Yasuda, M. Yamamoto, Validation of multiple single nucleotide variation calls by additional exome analysis with a semiconductor sequencer to supplement data of whole-genome sequencing of a human population, BMC Genomics 15 (2014) 673.
[6] J. Yasuda, K. Kinoshita, F. Katsuoka, I. Danjoh, M. Sakurai-Yageta, I.N. Motoike, Y. Kuroki, S. Saito, K. Kojima, M. Shirota, D. Saigusa, A. Otsuki, J. Kawashima, Y. Yamaguchi-Kabata, S. Tadaka, Y. Aoki, T. Mimori, K. Kumada, J. Inoue, S. Makino, M. Kuriki, N. Fuse, S. Koshiba, O. Tanabe, M. Nagasaki, G. Tamiya, R. Shimizu, T. Takai-Igarashi, S. Ogishima, A. Hozawa, S. Kuriyama, J. Sugawara, A. Tsuboi, H. Kiyomoto, T. Ishii, H. Tomita, N. Minegishi, Y. Suzuki, K. Suzuki, H. Kawame, H. Tanaka, Y. Taki, N. Yaegashi, S. Kure, F. Nagami, G. Tohoku, Medical Megabank Project Study, K. Kosaki, Y. Sutoh, T. Hachiya, A. Shimizu, M. Sasaki, M. Yamamoto, Genome analyses for the Tohoku Medical Megabank Project towards establishment of personalized healthcare, J. Biochem. 165 (2019) 139–158.
[7] Y. Shibuya, H. Tokunaga, S. Saito, K. Shimokawa, F. Katsuoka, L. Bin, K. Kojima, M. Nagasaki, M. Yamamoto, N. Yaegashi, J. Yasuda, Identification of somatic genetic alterations in ovarian clear cell carcinoma with next generation sequencing, Genes Chromosomes Cancer 57 (2018) 51–60.
[8] X. Pan, N. Nariai, N. Fukuhara, S. Saito, Y. Sato, F. Katsuoka, K. Kojima, Y. Kuroki, I. Danjoh, R. Saito, S. Hasegawa, Y. Okitsu, A. Kondo, Y. Onishi, F. Nagami, H. Kiyomoto, A. Hozawa, N. Fuse, M. Nagasaki, R. Shimizu, J. Yasuda, H. Harigae, M. Yamamoto, Monitoring of minimal residual disease in early T-cell precursor acute lymphoblastic leukaemia by next-generation sequencing, Br. J. Haematol. 176 (2017) 318–321.
[9] J.F. Boland, C.C. Chung, D. Roberson, J. Mitchell, X. Zhang, K.M. Im, J. He, S.J. Chanock, M. Yeager, M. Dean, The new sequencer on the block: comparison of Life Technology's Proton sequencer to an Illumina HiSeq for whole-exome sequencing, Hum. Genet. 132 (2013) 1153–1163.
[10] L.M. Bragg, G. Stone, M.K. Butler, P. Hugenholtz, G.W. Tyson, Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data, PLoS Comput. Biol. 9 (2013)e1003031.
[11] H. Seo, Y. Park, B.J. Min, M.E. Seo, J.H. Kim, Evaluation of exome variants using the Ion Proton Platform to sequence error-prone regions, PLoS One 12 (2017) e0181304.
[12] K.L. Huang, R.J. Mashl, Y. Wu, D.I. Ritter, J. Wang, C. Oh, M. Paczkowska, S. Reynolds, M.A. Wyczalkowski, N. Oak, A.D. Scott, M. Krassowski, A.D. Cherniack, K.E. Houlahan, R. Jayasinghe, L.B. Wang, D.C. Zhou, D. Liu, S. Cao, Y. W. Kim, A. Koire, J.F. McMichael, V. Hucthagowder, T.B. Kim, A. Hahn, C. Wang, M.D. McLellan, F. Al-Mulla, K.J. Johnson, N. Cancer Genome Atlas Research, O. Lichtarge, P.C. Boutros, B. Raphael, A.J. Lazar, W. Zhang, M.C. Wendl, R. Govindan, S. Jain, D. Wheeler, S. Kulkarni, J.F. Dipersio, J. Reimand, F. Meric-Bernstam, K. Chen, I. Shmulevich, S.E. Plon, F. Chen, L. Ding, Pathogenic germline variants in 10,389 adult cancers, Cell 173 (2018) 355–370 e14.
[13] D.A. Fruman, H. Chiu, B.D. Hopkins, S. Bagrodia, L.C. Cantley, R.T. Abraham, The PI3K pathway in human disease, Cell 170 (2017) 605–635.
[14] S. Koren, L. Reavie, J.P. Couto, D. De Silva, M.B. Stadler, T. Roloff, A. Britschgi, T. Eichlisberger, H. Kohler, O. Aina, R.D. Cardiff, M. Bentires-Alj, PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours, Nature 525 ( 2015) 114–118.
[15] D. Juric, I. Krop, R.K. Ramanathan, T.R. Wilson, J.A. Ware, S.M. Sanabria Bohorquez, H.M. Savage, D. Sampath, L. Salphati, R.S. Lin, H. Jin, H. Parmar, J.Y. Hsu, D.D. Von Hoff, J. Baselga, Phase I dose-escalation study of Taselisib, an oral PI3K inhibitor, in patients with advanced solid tumors, Cancer Discov. 7 (2017) 704–715.
[16] O. Reich, B. Liegl, K. Tamussino, S. Regauer, p185HER2 overexpression and HER2 oncogene amplification in recurrent vulvar Paget's disease, Mod. Pathol. 18 (2005) 354–357.
[17] R. Tanaka, Y. Sasajima, H. Tsuda, K. Namikawa, A. Tsutsumida, F. Otsuka, N. Yamazaki, Human epidermal growth factor receptor 2 protein overexpression and gene amplification in extramammary Paget disease, Br. J. Dermatol. 168 ( 2013) 1259–1266.
[18] O. Vornicova, D. Hershkovitz, T. Yablonski-Peretz, O. Ben-Itzhak, Z. Keidar, G. Bar-Sela, Treatment of metastatic extramammary Paget's disease associated with adnexal adenocarcinoma, with anti-HER2 drugs based on genomic alteration ERBB2 S310F, Oncologist 19 (2014) 1006–1007.
[19] J.S. Ross, K. Wang, L.M. Gay, R.N. Al-Rohil, T. Nazeer, C.E. Sheehan, T.A. Jennings, G.A. Otto, A. Donahue, J. He, G. Palmer, S. Ali, M. Nahas, G. Young, E. Labrecque, G. Frampton, R. Erlich, J.A. Curran, K. Brennan, S.R. Downing, R. Yelensky, D. Lipson, M. Hawryluk, V.A. Miller, P.J. Stephens, A high frequency of activating extracellular domain ERBB2 (HER2) mutation in micropapillary urothelial carcinoma, Clin. Cancer Res. 20 (2014) 68–75.
[20] G.S. Herter-Sprie, H. Greulich, K.K. Wong, Activating mutations in ERBB2 and their impact on diagnostics and treatment, Front. Oncol. 3 (2013) 86.
[21] M. Takata, N. Hatta, K. Takehara, Tumour cells of extramammary Paget's disease do not show either p53 mutation or allelic loss at several selected loci implicated in other cancers, Br. J. Cancer 76 (1997) 904–908.
[22] G. Giglia-Mari, A. Sarasin, TP53 mutations in human skin cancers, Hum. Mutat. 21 (2003) 217–228.

[23] T. Fujimura, S. Takahashi, T. Urano, J. Kumagai, T. Ogushi, K. Horie-Inoue, Y. Ouchi, T. Kitamura, M. Muramatsu, S. Inoue, Increased expression of estrogen-related receptor alpha (ERRalpha) is a negative prognostic predictor in human prostate cancer, Int. J. Cancer 120 (2007) 2325–2330.
[24] P. Sun, J. Sehouli, C. Denkert, A. Mustea, D. Konsgen, I. Koch, L. Wei, W. Lichtenegger, Expression of estrogen receptor-related receptors, a subfamily of orphan nuclear receptors, as new tumor biomarkers in ovarian cancer cells, J. Mol. Med. 83 (2005) 457–467.
[25] T. Suzuki, Y. Miki, T. Moriya, N. Shimada, T. Ishida, H. Hirakawa, N. Ohuchi, H. Sasano, Estrogen-related receptor alpha in human breast carcinoma as a potent prognostic factor, Cancer Res. 64 (2004) 4670–4676.
[26] I.S. Tam, V. Giguere, There and back again: the journey of the estrogen-related receptors in the cancer realm, J. Steroid Biochem. Mol. Biol. 157 (2016) 13–19.
[27] Y. Chudnovsky, D. Kim, S. Zheng, W.A. Whyte, M. Bansal, M.A. Bray, S. Gopal, M. A. Theisen, S. Bilodeau, P. Thiru, J. Muffat, O.H. Yilmaz, M. Mitalipova, K. Woolard, J. Lee, R. Nishimura, N. Sakata, H.A. Fine, A.E. Carpenter, S.J. Silver, R.G. Verhaak, A. Califano, R.A. Young, K.L. Ligon, I.K. Mellinghoff, D.E. Root, D.M. Sabatini, W.C. Hahn, M.G. Chheda, ZFHX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state, Cell Rep. 6 (2014) 313–324.
[28] T. Qing, S. Zhu, C. Suo, L. Zhang, Y. Zheng, L. Shi, Somatic mutations in ZFHX4 gene are associated with poor overall survival of Chinese esophageal squamous cell carcinoma patients, Sci. Rep. 7 (2017) 4951.
[29] M. Alhosin, Z. Omran, M.A. Zamzami, A.L. Al-Malki, H. Choudhry, M. Mousli, C. Bronner, Signalling pathways in UHRF1-dependent regulation of tumor suppressor genes in cancer, J. Exp. Clin. Cancer Res. 35 (2016) 174.
[30] C. Bronner, M. Achour, Y. Arima, T. Chataigneau, H. Saya, V.B. Schini-Kerth, The UHRF family: oncogenes that are drugable targets for cancer therapy in the near future? Pharmacol. Ther. 115 (2007) 419–434.
[31] S. Liu, H. Wang, L. Zhang, C. Tang, L. Jones, H. Ye, L. Ban, A. Wang, Z. Liu, F. Lou, D. Zhang, H. Sun, H. Dong, G. Zhang, Z. Dong, B. Guo, H. Yan, C. Yan, L. Wang, Z. Su, Y. Li, X.F. Huang, S.Y. Chen, T. Zhou, Rapid detection of genetic mutations in individual breast cancer patients by next-generation DNA sequencing, Hum. Genomics 9 (2015) 2.
[32] B.B. Tops, N. Normanno, H. Kurth, E. Amato, A. Mafficini, N. Rieber, D. Le Corre, A.M. Rachiglio, A. Reiman, O. Sheils, C. Noppen, L. Lacroix, I.A. Cree, A. Scarpa, M.J. Ligtenberg, P. Laurent-Puig, Development of a semi-conductor sequencing-based panel for genotyping of colon and lung cancer by the Onconetwork consortium, BMC Cancer 15 (2015) 26.
[33] L. Ding, M.H. Bailey, E. Porta-Pardo, V. Thorsson, A. Colaprico, D. Bertrand, D.L. Gibbs, A. Weerasinghe, K.L. Huang, C. Tokheim, I. Cortes-Ciriano, R. Jayasinghe, F. Chen, L. Yu, S. Sun, C. Olsen, J. Kim, A.M. Taylor, A.D. Cherniack, R. Akbani, C. Suphavilai, N. Nagarajan, J.M. Stuart, G.B. Mills, M.A. Wyczalkowski, B.G. Vincent, C.M. Hutter, J.C. Zenklusen, K.A. Hoadley, M.C. Wendl, L. Shmulevich, A.J. Lazar, D.A. Wheeler, G. Getz, N. Cancer genome atlas research, perspective on oncogenic processes at the end of the beginning of Cancer genomics, Cell 173 (2018) 305–320 e10.
[34] M.W. Lee, K.J. Jee, G.Y. Gong, J.H. Choi, K.C. Moon, J.K. Koh, Comparative genomic hybridization in extramammary Paget's disease, Br. J. Dermatol. 153 ( 2005) 290–294.
[35] E. De Braekeleer, N. Douet-Guilbert, F. Morel, M.J. Le Bris, A. Basinko, M. De Braekeleer, ETV6 fusion genes in hematological malignancies: a review, Leuk. Res. 36 (2012) 945–961.