OPEN

Citation: Human Genome Variation (2015) 2, 15030; doi:10.1038/hgv.2015.30

www.nature.com/hgv

DATA REPORT

Identification of novel exonic mobile element insertions in epithelial ovarian cancers

Takahide Hayano[1], Shiro Yamada[1], Kazuyoshi Hosomichi[1], Hirofumi Nakaoka[1], Kosuke Yoshihara[2], Sosuke Adachi[2], Katsunori Kashima[2], Kenichi Tanaka[2,3], Takayuki Enomoto[2] and Ituro Inoue[1]

Mobile elements comprise about half of the human genome. Three active mobile element families (L1, *Alu*, and SVA) possibly cause diseases such as cancer. We conducted mobile element insertion (MEI) profiling of 44 epithelial ovarian cancers using exome-sequencing data. We identified a total of 106 MEIs using the Mobster program, 8 of which were novel exonic MEIs.

Human Genome Variation (2015) **2,** 15030; doi:10.1038/hgv.2015.30; published online 1 October 2015

Mobile elements are major components of the human genome, comprising between 45 and 66% of the genome.[1–3] Mobile elements have contributed to the variation of the human genome.[3] Although most mobile elements have lost their mobile activity and have become fixed in the human genome, several mobile element families including L1, *Alu*, and SVA still retain their mobile activity. MEIs occasionally cause disease by disrupting gene functions, and >90 MEIs have been reported to cause disease, including breast cancer.[4] Recently, a large-scale MEI profiling of 11 human cancer types, including ovarian cancer, was performed as part of The Cancer Genome Atlas Pan-Cancer Project, and many novel MEIs were identified.[5] Here, we report the detection of MEIs in 44 Japanese patients with epithelial ovarian cancers (35 high-grade serous, 3 clear cell, 5 endometrioid, and 1 mixture of clear cell and endometrioid) using exome-sequencing data. We used an MEI detection algorithm, Mobster, which can be applied to exome-sequencing data.[6] PCR validation was performed for two types of PCR strategies. One type is MEI PCR, in which full-length MEIs are amplified using primers that bind to the outside of reference-MEI junctions. The other type is MEI-junction PCR, in which reference-MEI junctions are amplified using a primer map of the ME sequence (Supplementary Information 1A). We used touchdown PCR protocol (Supplementary Information 2). After PCR, expected amplicon sizes were checked by Bioanalyzer (Agilent, Santa Clara, CA, USA; Supplementary Information 1B).

The ethics committees of Niigata University and the National Institute of Genetics approved the study protocols, and each participant provided written informed consent for the collection of samples and subsequent analyses.

Genomic DNA was isolated from tumor tissues and the corresponding matched peripheral blood (normal tissue) from ovarian cancer patients. Then, genomic DNA was hybridized with SureSelect Human All Exon Kits (Agilent) and sequenced using Illumina HiSeq 2000 (Illumina, San Diego, CA, USA) with 90- or 100-base-paired end modules. The sequencing data were mapped to a human genome reference (hg19) using a standard method of BWA, Picard, and GATK, as previously described.[7]

MEI detection was conducted using Mobster with the default setting.[6] To manually review MEIs, MEIs with split read(s) were maintained. Mobster detected a total of 106 (29 L1 and 77 *Alu*) MEIs with split reads from the 44 ovarian cancer samples (Supplementary Information 3). We extracted exonic MEIs using bedtools.[8] Then, the regions previously reported as polymorphisms[5,6,9,10] with allelic fractions <0.005 (the number of split reads supporting MEI/the number of total reads spanning the insertion point) were removed.[5] Finally, eight exonic MEIs (3 L1 and 5 *Alu*) were identified as novel (Table 1). For the eight MEIs on eight genes (*OTUD7B*, *PDLIM7*, *RP1*, *XKR9*, *SLC30A8*, *KIFC2*, *OR1L4*, and *IPO4*), we manually reviewed the insertion points and zygosity with split reads by IGV (Figure 1a). We further performed PCR and were able to validate the existence of MEIs in seven of the eight genes (Table 1 and Supplementary Information 1). All seven MEIs were assumed to be heterozygous. Five of the seven MEIs were assumed to be somatic by comparison with matched normal data (Figure 1a and Supplementary Information 1). However, two of the five assumed somatic MEIs were also detected as germline MEIs in matched normal samples by PCR. These false MEI detections by Mobster in normal samples could be caused by insufficient sensitivity with low coverage of reads. When simulated exome data are used with homozygous MEIs, the sensitivity of Mobster is limited to the range from 52.7 to 85.4% at a depth of 10× to 160×, respectively.[6] The mean depth of our exome-sequencing data was 112×, ranging from 68× to 166×. For five of the seven MEIs, ME subfamilies were predicted using clipped sequences (Figure 1b) by CENSOR.[11] One ME (L1HS) belongs to an active subfamily, whereas the other four MEs (4 AluS) are members of ancient, inactive subfamilies.[12–14] Insertions of these ancient ME subfamilies may not occur through the canonical mechanism of target-primed reverse transcription,[3] but they can occur through genomic rearrangement, including chromosomal translocation.[15] In this study, we detected 106 MEIs from split reads of 44 samples (average 2.4, range 0–8 per sample). Fewer MEIs were detected compared with CEU trio's exome data set (average 7.3, range 6–8 per sample).[6] These results may suggest that some parameter adjustments of Mobster are required to detect more MEIs from the exome-sequencing data.

We found three novel somatic exonic MEIs for three genes (*PDLIM7*, *SLC30A8*, and *IPO4*) in ovarian cancer patients. Somatic

[1]Division of Human Genetics, National Institute of Genetics, Mishima, Japan; [2]Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan and [3]Niigata Medical Center Hospital, Niigata, Japan.
Correspondence: I Inoue (itinoue@nig.ac.jp)
Received 13 April 2015; revised 2 July 2015; accepted 4 July 2015

Table 1. Predicted novel exonic MEIs

Chr	*Insertion point*	*MEI*	*Sample*	*Histology*	*Gene*	*Zygosity*[a]	*PCR*	*Somatic/germline*[b]	*ME subfamily*[c]
1	149939291	L1	S004	Serous	*OTUD7B*	Heterozygous	Yes	Germline	Undetermined
5	176918034	ALU	S011	Serous	*PDLIM7*	Heterozygous	Yes	Somatic	AluSx
8	55540494	ALU	S271	Serous	*RP1*	Heterozygous	Yes	Germline	AluS
8	71593525	ALU	S005	Serous	*XKR9*	Heterozygous	Yes	Germline	Undetermined
8	118183382	L1	S031	Serous	*SLC30A8*	Heterozygous	Yes	Somatic	L1HS
8	145694721	ALU	S117	Serous	*KIFC2*	Heterozygous	Yes	Germline	AluS
9	125486418	L1	S068	Serous	*OR1L4*	Heterozygous	No	NA	NA
14	24655945	ALU	S008	Serous	*IPO4*	Heterozygous	Yes	Somatic	AluSc

Abbreviations: ME, mobile element; MEI, mobile element insertion; NA, not applicable. [a]Both zygosity and somatic status were manually assumed using IGV and experimentally validated by PCR. [b]Both zygosity and somatic status were manually assumed using IGV and experimentally validated by PCR. [c]MEI subfamilies were predicted using CENSOR. NA means not applicable.

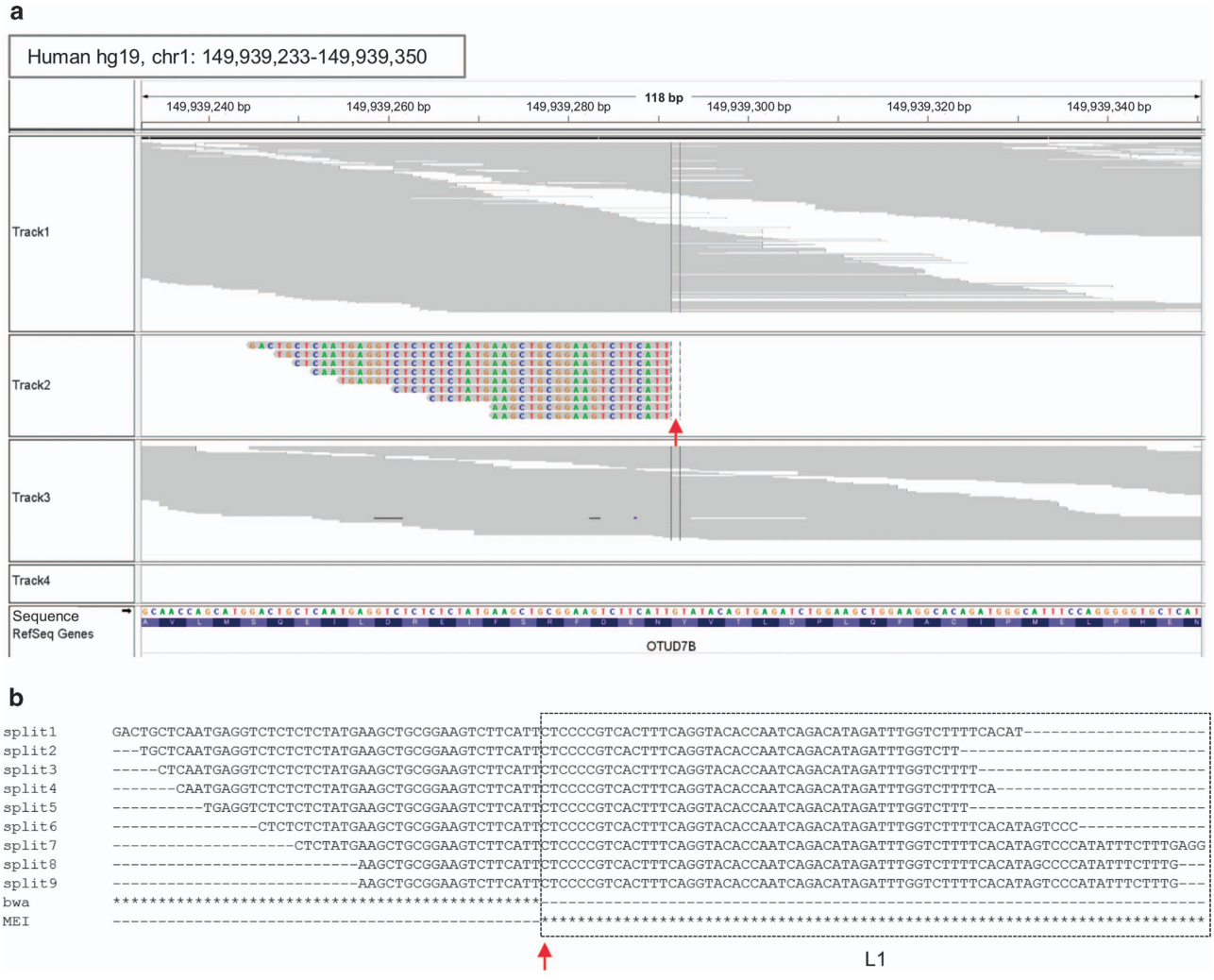

Figure 1. A representative novel exonic mobile element insertion (MEI). (**a**) Screenshot of IGV. Nine split reads (partially mapped to the reference genome and partially to a mobile element) were mapped on exon4 of *OTUD7B*. Track 1 shows mapped reads from tumors. Track 2 shows split reads from tumors extracted by Mobster. Track 3 shows mapped reads from matched normal tissue. Track 4 shows that no split reads were extracted from matched normal tissue by Mobster. The red arrow indicates the insertion point. (**b**) Alignment of the nine split reads. Asterisks in the BWA lane show nucleotides mapped to the reference genome (hg19). Asterisks in the MEI lane show clipped ME sequences (soft clipped by BWA). The red arrow indicates the insertion point.

exonic MEIs are rare events in cancer and have not previously been reported for ovarian cancer in previous studies.[5,10] PDLIM7, also known as LIM mineralization protein-1 (LMP-1), was reported to be a tumor suppressor in osteosarcoma cells.[16] IPO4, also known as Importin 4, has a role in the nuclear import of FANCD2, which is associated with tumor suppression.[17] Disruptions of PDLIM7 and IPO4 may be associated with ovarian cancer malignancy and a deficiency in a tumor-suppressing mechanism, respectively. A nonsense mutation in *SLC30A8* was found in non-triple-negative breast cancer.[18] We found four germline exonic

MEIs in four genes (*OTUD7B*, *RP1*, *XKR9*, and *KIFC2*). OTUD7B was reported to control the non-canonical NF-kB pathway, the activation of which is important for cell growth and survival.[19] XKR9 enhanced apoptosis by promoting phosphatidylserine exposure.[20] Disruption of OTUD7B and XKR9 would be advantageous to ovarian cancer cell growth and survival. The functional relevance of RP1 and KIFC2 to cancer has not yet been reported. Further studies are required to investigate the functional significance of these gene products for ovarian cancer. We were not able to detect an MEI for OR1L4 by PCR. The relatively low allelic fraction of this MEI (Supplementary Information 4) suggests that it is a false positive.

In conclusion, we identified 106 MEIs from 44 ovarian cancer samples, including 8 novel exonic MEIs. PCR validation demonstrated a high specificity of exonic MEI detection (87.5%, 7/8). This is the first report of MEI detection in ovarian cancer using exome-sequencing data. This approach will provide new insights for studies of cancer-causing mutations.

HGV DATABASE

The relevant data from this Data Report are hosted at the Human Genome Variation Database at http://dx.doi.org/10.6084/m9.figshare.hgv.672, http://dx.doi.org/10.6084/m9.figshare.hgv.674, http://dx.doi.org/10.6084/m9.figshare.hgv.676, http://dx.doi.org/10.6084/m9.figshare.hgv.678, http://dx.doi.org/10.6084/m9.figshare.hgv.680, http://dx.doi.org/10.6084/m9.figshare.hgv.682, http://dx.doi.org/10.6084/m9.figshare.hgv.684, http://dx.doi.org/10.6084/m9.figshare.hgv.686.

COMPETING INTEREST

The authors declare no conflict of interest.

REFERENCES

1 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J *et al.* Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860–921.
2 de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 2011; **7**: e1002384.
3 Xing J, Witherspoon DJ, Jorde LB. Mobile element biology: new possibilities with high-throughput sequencing. *Trends Genet* 2013; **29**: 280–289.
4 Hancks DC, Kazazian HH Jr. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* 2012; **22**: 191–203.
5 Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res* 2014; **24**: 1053–1063.
6 Thung D, de Ligt J, Vissers L, Steehouwer M, Kroon M, de Vries P *et al.* Mobster: accurate detection of mobile element insertions in next generation sequencing data. *Genome Biol* 2014; **15**: 488.
7 Hayano T, Yokota Y, Hosomichi K, Nakaoka H, Yoshihara K, Adachi S *et al.* Molecular characterization of an intact p53 pathway subtype in high-grade serous ovarian cancer. *PLoS ONE* 2014; **9**: e114491.
8 Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; **26**: 841–842.
9 Stewart C, Kural D, Strömberg MP, Walker JA, Konkel MK, Stütz AM *et al.* A comprehensive map of mobile element insertion polymorphisms in humans. *PLoS Genet* 2011; **7**: e1002236.
10 Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ 3rd *et al.* Landscape of somatic retrotransposition in human cancers. *Science* 2012; **337**: 967–971.
11 Kohany O, Gentles AJ, Hankus L, Jurka J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* 2006; **7**: 474.
12 Mir AA, Philippe C, Cristofari G. euL1db: the European database of L1HS retrotransposon insertions in humans. *Nucleic Acids Res* 2015; **43**: D43–D47.
13 Wagstaff BJ, Kroutter EN, Derbes RS, Belancio VP, Roy-Engel AM. Molecular reconstruction of extinct LINE-1 elements and their interaction with non-autonomous elements. *Mol Biol Evol* 2013; **30**: 88–99.
14 Batzer MA, Deininger PL, Hellmann-Blumberg U, Jurka J, Labuda D, Rubin CM *et al.* Standardized nomenclature for Alu repeats. *J Mol Evol* 1996; **42**: 3–6.
15 Bunting SF, Nussenzweig A. End-joining, translocations and cancer. *Nat Rev Cancer* 2013; **13**: 443–454.
16 Liu H, Huang L, Zhang Z, Zhang Z, Yu Z, Chen X *et al.* LIM mineralization protein-1 inhibits the malignant phenotypes of human osteosarcoma cells. *Int J Mol Sci* 2014; **15**: 7037–7048.
17 Wang J, Sarkar TR, Zhou M, Sharan S, Ritt DA, Veenstra TD *et al.* CCAAT/enhancer binding protein delta (C/EBPdelta, CEBPD)-mediated nuclear import of FANCD2 by IPO4 augments cellular response to DNA damage. *Proc Natl Acad Sci USA* 2010; **107**: 16131–16136.
18 Horvath A, Pakala SB, Mudvari P, Reddy SD, Ohshiro K, Casimiro S *et al.* Novel insights into breast cancer genetic variance through RNA sequencing. *Sci Rep* 2013; **3**: 2256.
19 Hu H, Brittain GC, Chang JH, Puebla-Osorio N, Jin J, Zal A *et al.* OTUD7B controls non-canonical NF-κB activation through deubiquitination of TRAF3. *Nature* 2013; **494**: 371–374.
20 Suzuki J, Imanishi E, Nagata S. Exposure of phosphatidylserine by Xk-related protein family members during apoptosis. *J Biol Chem* 2014; **289**: 30257–30267.

Supplemental Information for this article can be found on the *Human Genome Variation* website (http://www.nature.com/hgv).