LINE-1 retrotransposon-mediated DNA transductions in endometriosis associated ovarian cancers

Zhouchunyang Xia \textsuperscript{a}, Dawn R. Cochrane \textsuperscript{b}, Michael S. Anglesio \textsuperscript{a,c}, Yi Kan Wang \textsuperscript{b}, Tayyebeh Nazeran \textsuperscript{a}, Basile Tessier-Cloutier \textsuperscript{a}, Melissa K. McConney \textsuperscript{d}, Janine Senz \textsuperscript{b}, Amy Lum \textsuperscript{b}, Ali Bashashati \textsuperscript{b}, Sohrab P. Shah \textsuperscript{a,b}, David G. Huntsman \textsuperscript{a,b,*}

\textsuperscript{a} Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada
\textsuperscript{b} Department of Molecular Oncology, BC Cancer Agency, Vancouver, BC, Canada
\textsuperscript{c} Department of Gynecology and Obstetrics, University of British Columbia, Vancouver, BC, Canada
\textsuperscript{d} Department of Human Genetics, McGill University, Research Institute of the McGill University Health Centre, Montreal, QC, Canada

HIGHLIGHTS

- An active LINE-1 retrotransposon from TTC28 gene is present in 32% of EAOCs.
- TTC28 LINE-1 events are near ubiquitous across multiple EAOC tumor samplings.
- TTC28 LINE-1 is likely activated early during EAOC development.
- TTC28 LINE-1 may be used to track EAOC development.

ABSTRACT

Objective. Endometrioid (ENOC) and clear cell ovarian carcinoma (CCOC) share a common precursor lesion, endometriosis, hence the designation endometriosis associated ovarian cancers (EAOC). Long interspersed nuclear element 1 (LINE-1 or L1), is a family of mobile genetic elements activated in many cancers capable of moving neighboring DNA through 3’ transductions. Here we investigated the involvement of specific L1-mediated transductions in EAOCs.

Methods. Through whole genome sequencing, we identified active L1-mediated transductions originating within the TTC28 gene in 34% (10/29) of ENOC and 31% (11/35) of CCOC cases. We used PCR and capillary sequencing to assess the presence of specific TTC28-L1 transductions in formalin-fixed paraffin-embedded (FFPE) blocks from six different anatomical sites (five tumors and one normal control) for four ENOC and three CCOC cases, and compared the results to the presence of single nucleotide variations (SNVs)/frame shift (fs) mutations detected using multiplex PCR and next generation sequencing.

Results. TTC28-L1 mediated transductions were identified in at least three tumor samplings in all cases, and were present in all five tumor samplings in 5/7 (71%) cases. In these cases, KRAS, PIK3CA, CTNNB1, ARID1A, and PTEN mutations were found across all tumor sites while other selected SNV/fs mutations of unknown significance were present at varying allelic frequencies.

Conclusion. The TTC28-L1 transductions along with classical driver mutations were near ubiquitous across the tumors, suggesting that L1 activation likely occurred early in the development of EAOCs. TTC28-L1 transductions could potentially be used to determine clonal relationships and to track ovarian cancer progression.

© 2017 Published by Elsevier Inc.

1. Introduction

Ovarian cancer is the most lethal gynecological cancer, and it is the 5th leading cause of cancer related deaths in women in the United States [1]. Endometrioid ovarian carcinoma (ENOCC) and clear cell ovarian carcinoma (CCOC) each account for approximately 10% of all ovarian cancer cases.
cancer cases [2]. Both cancers are low grade subtypes that are often di-
agnosed at stages I and II. When diagnosed at stages III/IV, ENOC has a
more favorable prognosis compared to CCOC, as CCOC tends to have a
much lower response rate to chemotherapy [2,3]. A recent study of
the genomic profile of ovarian cancers stratified ENOC into three sub-
groups: ultramutator (high mutation load), micro-satellite instable (MSI),
and micro-satellite stable (MSS), and CCOC cases into two sub-
groups: APOBEC (localized hypermutation) and mutation signatures
similar to one previously associated with age, though not age-related
in CCOC [4].

Despite having different pathology and genomic profiles, both ENOC
and CCOC are associated with endometriosis, hence the designation en-
dometriosis associated ovarian cancer (EAOC). Endometriosis is char-
acterized by ectopic growth of endometrial epithelium and stroma.
Women with endometriosis have a 2- to 3-fold increased risk of devel-
oping ENOC and CCOC [5]. Common genes with mutations implicated to
deriving ENOC and CCOC [5]. Common genes with mutations implicated to
drive tumorigenesis include ARID1A (30%), PTEN (20%), and CINN1B
(38–50%) in ENOC, and ARID1A (46–57%) and PIK3CA (33–46%) in
CCOC. Cancer driving mutations in ARID1A, PTEN, and PIK3CA have also
been found in endometriotic lesions both adjacent and distant to the tu-
mors, indicating a clonal relationship between these diseases [6–9]. In
addition, canonical cancer mutations have been found in deep infiltrat-
ing endometriosis that is not associated with cancer [7]. Given that driv-
er mutations in EAOCs are only found in a subset of patients, and certain
changes can occur in endometriosis without cancer, additional types of
biomarkers are needed to distinguish women with endometriosis at risk
of developing cancer.

In our whole genome sequencing of 29 ENOC and 35 CCOC cases, we
observed frequent rearrangement events originating from an active
LINE-1 (L1) retrotransposon located in the 7TC28 gene on chromosome
22q12. Long interspersed nuclear elements 1 (LINE-1 or L1), are rer-
itive mobile DNA elements that accounts for 1% of the human genome
[10–13]. L1s encode an endonuclease and a reverse transcriptase that allows them to “copy-and-paste” their own sequences into random ge-
nomic locations via RNA intermediates [10,12]. They are also capable of
taking unique downstream DNA fragments with them during this ge-
netic propagation in a process called 3’ transduction [10]. While most
L1s are not functional due to structural rearrangements, there are ap-
proximately 80–100 full-length L1s that are potentially active [10]. L1
activation is important in the germ cells during embryogenesis [14],
but epigenetic silencing typically occurs in adult [12]. In the cancer ge-
nomes, however, L1s become activated as part of a general degradation
of genomic and epigenetic integrity [12].

The effects of L1 retrotransposon activation in cancer development
and progression has been studied in various cancers, including
head and neck, prostate, ovarian, colorectal, gastric, esophageal, pan-
creatic, and breast cancer [10,11,13,15–21]. While most studies sur-
veyed the landscape of somatic L1 retrotranspositions by detecting the
presence of novel inserted L1 sequences in tumors, Tubio and collea-
gues studied the effect of 3’ transduction events mediated by a set of active L1s in cancers [10]. This specific 7TC28-L1-
mediated transduction from chromosome 22q12 is one of the most
frequently active L1s found in colon, head and neck, and lung cancer
samples and is the only active L1 element in 93% of breast cancer
samples [10]. Specific L1-mediated transductions have been less ex-
tensively studied in EAOCs [17,21].

Here we used PCR and capillary sequencing to validate the presence of
7TC28-L1-mediated transductions in different tumor samplings of
ENOC and CCOC, and we compared the results to targeted re-
sequencing of selected single nucleotide variations (SNVs)/frame shift
(fs) mutations in the same tissues to infer biological timing of L1 tran-
ductions. Our data showed that these transduction events appeared with canonical driver mutations for majority of the cases, and suggests that they likely occurred early in tumorigenesis. We hope to use
7TC28-L1 mediated DNA transductions to gain insights into the tumor
development of EAOCs.

2. Methods

2.1. Whole genome sequencing (WGS)

The cohort and methods have been described previously by Wang
et al. [4]. Briefly, tumor (frozen tissue) and matched normal (buffy
cast) DNA libraries were constructed for 29 ENOC and 35 CCOC cases,
and sequenced using Illumina HiSeq 2500 V4 chemistry. Analyses
were performed using bioinformatic methods described in Wang et al.
 supplemental materials [4].

2.2. Case selection

Four ENOC and three CCOC cases with the highest numbers of
7TC28-L1-mediated transductions detected by WGS were selected from
our previous study [4]. All materials were provided by the
OV Care gynecological tissue bank. H&E slides for all available
formalin-fixed paraffin-embedded (FFPE) blocks were reviewed by ex-
pert pathologists (T.N. and B.T-C.). Five FFPE tumor tissue blocks with
the highest cellularity (>90%) and one representative FFPE normal tis-
ue block were selected for each of the seven cases. Overall project pro-
cesses were approved by the BC Cancer Agency or the University of
British Columbia Research Ethics Board (REB #H08-01411 and #H09-
02153).

2.3. DNA extraction from FFPE blocks

FFPE tissue blocks were cut 10um thick and applied on charged glass
slides, where relevant areas, identified by a pathologist (T.N.), were
macrodissected. DNA was extracted from FFPE tissues using the QiAamp
dNA FFPE Tissue Kit (Qiagen) as per manufacturers’ protocols. All DNA
was quantified using the broad range DNA assay Qubit fluorimeter (Life
Technologies).

2.4. PCR validation of 7TC28-L1-mediated transduction events

To validate all L1 transduction events, primers were designed using
Primer3 to generate 170–250 bp amplicons, which spanned the trans-
duction junction. Primers were validated on WGA (whole genome ampl-
ified) DNA extracted from frozen tumor tissues (positive control) and
buffy coats (negative control) for each of the cases. Primers were tested
at an annealing temperature of 60 °C, using PCR SuperMix High Fidelity
mastermix (ThermoFisher). Successful PCR was indicated by a single
band at 200–300 bp using gel electrophoresis. For failed PCRs (no
band, a smear, or multiple bands), a gradient PCR from 55 °C to 64 °C
was performed. If the PCRs were still unsuccessful, Platinum® Taq
dNA Polymerase High Fidelity (ThermoFisher) were used in substitu-
tion. Successful PCRs were validated via sequencing using ABI 3130XL
automated capillary sequencing as per manufacture protocol. Validated
primers were then used on the FFPE-extracted DNA samples with 50 ng
input.

2.5. Microfluidic PCR validation of selected SNV/Frameshift mutations

From the whole genome sequencing results, six likely pathogenic
missequences and/or frameshift mutations were selected for each case for
validations on the FFPE tissue blocks. Primer sets were designed using
Primer 3 to amplify the specific gene regions. Forward primers were
tagged with CS1 (5’-ACATCGACGCTAGTCTTACA-3’) and reverse
primers with CS2 (5’-TACGGTAGACAGTGCTTCTCC-3’) sequencing
tags. PCR products (150–200 bp) were amplified using the Fluidigm
48 × 48 Access Arrays, as per manufacturers’ protocol, with input of
50 ng FFPE derived DNA. DNA barcodes (10 bp) with Illumina cluster-
generating adapters were added to the libraries post-Fluidigm harvest,
and cleaned-up using Agencourt AMPure XP beads (Beckman Coulter).
Barcoded PCR products were then quantified using the high sensitivity

Please cite this article as: Z. Xia et al., LINE-1 retrotransposon-mediated DNA transductions in endometriosis associated ovarian cancers, Gynecol Oncol (2017), https://doi.org/10.1016/j.ygyno.2017.09.032
DNA assay Qubit fluorometer (Life Technologies) and pooled to one total library by normalizing to equal amounts of PCR product. In total, 48 samples were pooled and denatured according to Illumina standard protocols, and sequenced using a MiSeq 300 cycle V2 kit on the Illumina MiSeq for ultra-deep validations. Uni-directional barcode sequencing was performed. Analysis was performed using the bam and VCF files generated through Illumina MiSeq reporter.

3. Results

3.1. TTC28-L1 mediated transductions are frequent events in clear cell and endometrioid ovarian cancer

Whole genome shotgun sequencing of 29 ENOC and 35 CCOC cases to a median coverage of 51× and 37× for the tumor and matched normal DNA respectively was performed in a previous study [4]. Here, we studied this same data set generated via a structural variant caller (deStruct [22]), and found that the only recurrent rearrangement events originated from the TTC28 gene at the chromosome 22q12 locus (Fig. 1). Previously thought to be a fragile site for translocation events, we now know that this is a L1-retrotransposition event, as indicated by the clustering of breakpoints within 1 kb downstream of an active L1. This retrotransposon-mediated transduction event was observed in 34% (10/29) of ENOC, and 31% (11/35) of CCOC cases. None of the transductions targeted exons. There was no association between TTC28-L1 mediated transductions and common mutations or landscape features (data not shown). Because L1 elements are highly repetitive throughout the genome, there is difficulty aligning the sequences. Consequently, there could be additional L1 retrotransposition events that were undetected.

3.2. TTC28-L1 mediated transductions are early events in ENOC and CCOC oncogenesis

We validated these TTC28-L1 transductions on high quality gnomic DNA extracted from patient frozen tumors and buffy coats via PCR, with primers designed to flank the sequences around the insertion breakpoint, such that only samples with the transductions amplified. Amplified PCR products are sequence on the Sanger to confirm the presence of breakpoints with DNA sequences from both sides. While the TTC28-L1 transductions that failed at the PCR or Sanger validation stages were excluded, each case had at least one validated transduction event (for a list of validated TTC28-L1 transductions, see Supplemental Table S1). As expected, validated TTC28-L1 transduction events were found only in the tumor and not the buffy coat samples.

To infer the cellular timing of these TTC28-L1-mediated transductions, we took archival FFPE tumor samplings from five different sites for each of the seven patients. We hypothesized that the same transduction events detected at multiple tumor sites should indicate early activation. For each case, we selected six somatic SNVs/fs mutations detected via WGS and with highest frequencies and predicted impact, and performed targeted re-sequencing in the tumor samplings to further characterize the timing of the TTC28-L1 mediated transductions relative to WGS detected mutations (for a detailed list of SNVs/fs, see Supplemental Table S2).

We detected TTC28-L1 insertions in all FFPE tumor sites in the majority of the cases (5/7 cases, ENOC 1–4 and CCOC 1) (Fig. 2). In these cases, SNVs and fs mutations were detected to varying allelic frequencies for each tumor block, reflecting both inter- and intra-tumor heterogeneity (Fig. 2 and Supplemental Table S2). For each case, SNVs/fs mutations that were found in every tumor block sampled tend to be canonical cancer driver mutations, such as mutations in ARID1A, PIK3CA, KRAS, PTEN.
Fig. 2. (a) TTC28-L1 retrotransposition presence and allelic frequencies for five selected SNV/ins mutations at six anatomical sites for each of the seven ENOC and CCOC cases. Three shades of grey portray differences in allelic frequencies: dark grey indicates high frequency (≥50%) at the point mutation; lighter grey indicates a frequency between 20% and 50%; lightest grey indicates low frequency (5%–20%). Red box indicates the presence of TTC28-L1 retrotransposition, and white box indicates no variance or retrotransposition detected. Numbers indicate the allelic frequencies in percentage. A.F., allelic frequencies; Ov, ovary; Cer, cervix; Ut, uterus; R, right; L, left; N, normal; ND, not detected. (b) Circos plot displays retrotransposition events originating from TTC28 (Chr. 22q12) for the seven selected cases. All retrotransposition targets were in non-coding spaces. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: Z. Xia, et al., LINE-1 retrotransposon-mediated DNA transductions in endometriosis associated ovarian cancers, Gynecol Oncol (2017), https://doi.org/10.1016/j.ygyno.2017.09.032
and CTNN1B (Fig. 2 and Supplemental Table S2). These results suggest that TTC28-L1 was likely activated early during tumorigenesis, and was mediating DNA transductions along with SNVs/fs mutations in the earlier tumor clones. In one of the five cases (ENOC 4) the TTC28-L1 transduction was also detected in the histologically normal cervical tissue block used as a normal FFPE control, while the insertion event was not found in the same patient’s buffy coat DNA. A second sampling from this tissue block yielded the same result, which may indicate the presence of histologically normal cells clonally related to the tumor, and suggest that L1 activation could occur even earlier.

In a single CCOC case (CCOC 2), we observed what appears to be two distinct clonal populations within the same ovarian tumor, where three of the five tumor blocks surveyed had highly similar SNVs/fs allelic frequencies in addition to presence of the TTC28-L1 transduction event, while the other two tumor blocks contained neither SNVs/fs nor the TTC28-L1 transduction events.

Four cases in our cohort (ENOC 1, ENOC 2, ENOC4, and CCOC3) had synchronous tumors in the ovary and the uterus. In the three endometrioid ovarian carcinoma cases (ENOC 1, ENOC2, and ENOC 4), we observed the TTC28-L1 transduction event as well as several shared SNV mutations in both the ovarian and uterine tumors. The observation of the same TTC28-L1 transductions in tumors at anatomically different locations further suggests that L1 activation occurred early in these diseases. In the clear cell ovarian carcinoma case (CCOC 3), SNVs/fs mutations and the TTC28-L1 event were found in all the samplings of the ovarian tumor but were not seen in uterine tumor.

4. Discussion

Herein we described the presence of DNA transductions mediated by a specific and active retrotransposon, TTC28-L1, in the endometriosis associated ovarian carcinoma subtypes: ENOC and CCOCC. Our data showed that TTC28-L1 mediated transductions was detectable across all tumor sites in the majority of the cases. The presence of TTC28-L1 transduction was sometimes but not always accompanied by SNV/frameshift mutations at various allelic frequencies. It is highly unlikely that the same retrotransposition event arose independently at every tumor site, which indicates that TTC28-L1 transductions were not subclonal events, but more likely events that occurred early in tumorigenesis. In addition, the presence of TTC28-L1 transductions correlated with high allelic frequencies of driver mutations implicated in the development of these cancers, including mutations in ARID1A, CTNNB1, PIK3CA, and PTEN [8], supporting our hypothesis that TTC28-L1 retrotransposons are activated early in EAOV development.

Targeted resequencing was previously used by our group to describe a clonal relationship between the endometrial and ovarian cancers in three of the four synchronous endometrioid and ovarian tumors, ENOC 2, ENOC 4 and CCOC 3 [23]. We also found somatic mutations and TTC28-L1 events shared between the uterine and ovarian tumors in ENOC 2 and ENOC 4, however, we did not detect any of our selected mutations nor the TTC28-L1 event in the uterine tumor from CCOC 3. CCOC 3 was unusual in several ways: different pathologies observed in the ovary (clear cell) and uterus (endometrioid), and there was only one single shared somatic event (a different event from the ones we repeated) as reported by Anglesio et al. [23]. It is possible this somatic event had occurred before either the seeding of endometriosis or the activation of the TTC28-L1 retrotransposon.

It is interesting to note that 3’ transductions accounts for approximately one quarter of all somatic L1 retrotranspositions, and the majority of L1 retrotranspositions result in solo L1 insertions, either as a full L1 or truncated at the 5’ end [10]. This means that there are likely more L1 insertions undetected in our cohort. Past studies on somatic L1 retrotranspositions in various cancer types reported much more solo L1 insertions than 3’ transduction events; however, the lack of L1 3’ transductions detected could be due to bias in L1 sequencing and analysis methods [16,18–20]. Nonetheless, we have found TTC28-L1 transduction events at a high frequency in our cohort, which could be a unique feature of EAOCs.

While somatic L1 insertions are common in certain tumor types, there is little evidence that they are oncogenic. Only two studies showed that somatic L1 retrotranspositions (solo L1 insertions) directly initiated oncogenesis in colorectal cancer via a somatic L1 insertion in an exon of the tumor suppressor gene APC [24,25]. The majority of somatic L1 insertions target heterochromatic regions with no obvious functional impact [10,12]. We also observe that our L1 transductions targeted non-coding regions within out cohort.

Most research in the field focuses on using L1 activations as a marker of tumor development, by assessing the global hypomethylation of L1 loci and the expression of L1 mRNA and proteins throughout stages of cancer development and metastasis. The general conclusion is that somatic L1 retrotranspositions are passenger events that occur after a permissive environment has been established during cancer development, such as dysregulated epigenetic control [10,12,17]. Evidence of L1 activation as an early event include a recent study that found stepwise decrease of methylation across different L1 loci have been observed between normal endometrium, contiguous endometriosis (endometriosis adjacent to tumor), and ENOCs/CCOCs [17]. In gastrointestinal (GI) cancers, shared somatic L1 retrotranspositions (solo insertions) have been found between precancerous lesions and colon tumors [20], as well as between Barrett’s esophagus (precursor lesion of esophageal cancer) and concomitant esophageal tumors [19], suggesting early activation of L1 during GI tumor development. This is likely the scenario in our cohorts, where the TTC28-L1 was activated early during ENOC/CCOC tumorigenesis, but after a retrotransposition-permissive environment was established.

The exact relationship between TTC28-L1 activation and malignant transformation in these two ovarian cancer subtypes is challenging to elucidate. Further studies will be needed to assay endometriosis tissues that are adjacent to tumors, distant from tumors, as well as endometriosis without cancer involvement. Such studies could determine whether TTC28-L1 transduction is a useful marker for cancer risk of endometriosis.

EAOCs suffer from a lack of effective early detection tools. TTC28-L1 and other L1s can potentially be used as biomarkers to identify early malignant transformations, especially in the subset of CCOC/ENOC cases where common coding mutations are not present. In addition, a better understanding of L1 activation and transduction patterns over the course of ENOC and CCOC progression may help elucidate the underlying mechanisms of ovarian carcinogenesis and progression.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jygyno.2017.09.032.

Conflict of interest statement
Dr. David Huntsman reports that he is a founder and CMO of Contextual Genomics, a cancer testing company; however, Contextual does not test for the features described in this manuscript. Otherwise the authors have no conflict of interest.

Acknowledgement

The authors would like to thank the members of the Huntsman group and the OVCARE team for all their technical supports. This research would not have been made possible without the gracious donation of tissues from the women with ovarian cancer. This study was supported by grants from the Canadian Cancer Society (Impact Grant # 701603), BC Cancer Foundation and the VGH & UBC Hospitals Foundations (to OVCARE; http://www.ovcare.ca).

References

Please cite this article as: Z. Xia, et al., LINE-1 retrotransposon-mediated DNA transductions in endometriosis associated ovarian cancers, Gynecol Oncol (2017), https://doi.org/10.1016/j.ygyno.2017.09.032