Sex steroids regulate genes containing intragenic LINE-1 during menstrual cycles

Saichon Chaiwongwatanakul[1], Viroj Boonyaratanakornkit[2], Sissades Thongsima[3], Apiwat Mutirangura[4], Pattamawadee Yanatatsaneejit[5]*

[1]Department of Biological Sciences, Faculty of Science, Chulalongkorn University, Bangkok, Thailand
[2]Department of Clinical Chemistry, Faculty of Allied Health Sciences, Chulalongkorn University, Bangkok, Thailand
[3]Biostatistics and Informatics Laboratory, Genome Institute, National Center for Genetic Engineering and Biotechnology, Thailand Science Park, Pathumtani, Thailand
[4]Center of Excellence in Molecular Genetics of Cancer and Human Diseases, Department of Anatomy, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand
[5]Human Genetics Research Group, Department of Botany, Faculty of Sciences, Chulalongkorn University, Bangkok, Thailand.

**Corresponding author: yanpatta@hotmail.com*

ABSTRACT

Long interspersed element-1s (LINE-1s) are retrotransposable elements. Recent studies have shown that genes containing LINE-1s are regulated by epigenetic regulation of intragenic LINE-1s in cancer and in embryonic development. Moreover, intragenic LINE-1 may possess a regulatory function under physiological conditions. We hypothesized that female sex steroids could alter expressions of genes containing LINE-1. Here we analyzed the influence of intragenic LINE-1s on gene expression during menstrual phases by using expression array data from public database (Gene Expression Omnibus (GEO), (http://www.ncbi.nlm.nih.gov/geo). Gene expression levels in early, mid and late secretory phases were compared with those in the proliferative phase. Genes containing intragenic LINE-1s were highly and significantly up-regulated in early-secretory phase as compared to those in the proliferative phase in both GSE tests (GSE6364 with p< 1.05E-10; OR = 1.62; 95% CI = 1.40-1.88 and GSE4888 with p<2.22E-10; OR=1.61; 95% CI=1.39-1.87). Genes containing intragenic LINE-1 were significantly up-regulated in both mid- and late-secretory phase as compared to those in the proliferative phase (GSE4888 with p< 5.66E-03; OR=1.24; 95% CI = 1.06-1.44 and p<1.14E-04; OR=1.42; 95% CI=1.19-1.69). However, genes containing intragenic LINE-1 were found to be significantly down-regulated only in mid-secretory phase compared to those in the proliferative phase. (GSE6364 p<2.98E-02; OR=1.15; 95% CI=1.01-1.31 and GSE4888 p<6.22E-04; OR=1.25; 95% CI=1.10-1.41). Together, these data suggest that intragenic LINE-1 could play roles in regulation and expression of sex steroid regulated genes during menstrual cycles.

Keywords: estrogen, progesterone, menstrual phases, genes containing LINE-1

INTRODUCTION

The menstruation cycle is the important mechanism in female mammals where cyclic changes prepare the uterine lining for fertilized egg. In humans, the menstruation cycle in the uterus composes of proliferative (follicular) and secretory (luteal) phases which are regulated by sex hormones. These sex hormones include estrogen (E_2) and progesterone which play significant roles in changing the uterus during the cycle. Critically high levels of E_2 in the proliferative phase (PR) activate ovulation by inducing follicle-stimulating hormone (FSH) and luteinizing hormone (LH) secretion. The lower level of E_2 in the late PR then causes the lowering of FSH and LH levels. In secretory phase (SE), influence of E_2 and progesterone secretion can continue to build up the endometrium in preparation for implantation. Degeneration of corpus luteum and the beginning of menstruation will present when levels of these two hormones decline (fertilization is not complete). Hormonal levels in the menstruation cycle are very important mechanisms as they have influential organs in the reproductive system. Controlling gene expression is also related to complex interactions, such as signal transduction pathways, general and specific transcription factors, chromatin remodeling proteins and even the methylation of intragenic repetitive sequences (Ahmad and Kumar, 2011;

Aporntewan *et al.*, 2011). Up- or down- regulation of genes can be induced by various factors. Sex hormones can initiate up-regulated of target genes by acting as transcription factors after binding to specific receptors, meanwhile their receptor agonists may cause down-regulated effects of their respective receptors (Wierman, 2007).

Long interspersed nuclear element-1 (LINE-1 or L1) is a family of non-LTR retrotransposons. It presents about 18–20% with more than 500,000 copies of repetitive sequences in the human genome (Lander *et al*., 2001). These repetitive sequences disperse all over the chromosomes and affect the regulation of human host gene transcription (Matlik *et al*., 2005). Previously, we demonstrated that methylation of intragenic full length LINE-1 is an epigenetic modification that regulates gene expression in cancer (Aporntewan *et al.*, 2011) and embryogenesis (Ngamphiw *et al.*, 2014) . Moreover, our previous studiesfound that genes containing LINE-1s in neutrophils from the systemic lupus erythematosus (SLE) patients showed significantly up-regulated compared with the healthy controls (Sukapan *et al.*, 2014). We also found that genes containing LINE-1s expression can be altered by a number of siRNA experiments, including siRNA to estrogen receptors (W anichnopparat *et al.*, 2013). Therefore, we hypothesized that intragenic LINE-1 possess a *cis* regulatory function and alters genes containing LINE-1 expression during menstrual cycles.

Base on microarray data of the gene expression study, it can provide the interesting information regarding the comparison of genes containing the LINE-1 library. The screening of gene expression and the lists of genes containing LINE-1 which are related to each phase of menstruation cycle were done. Recently, there are no reports and studies which explain the relationship between each phase in themenstruation cycle and genes containing LINE-1 expression. This is the first study to provide more function of LINE-1 related to sex hormone levels in human gene expression control. Our study aims to verify the relationship between genes containing LINE-1 expression of endometrium in the menstruation cycle which presents the different dominant hormonal level in humans.

MATERIALS AND METHODS

Expression of genes containing LINE-1 and mRNA classification

Intragenic LINE-1s (from L1 base, Penzkofer *et al*., 2005) were categorized according to their genomic locations in the NCBI Reference Sequence (RefSeq) annotation which was the same as previously reported (Aporntewan *et al.*, 2011). To evaluate the influence of LINE-1s to host gene expression, genes were divided into two categories, containing LINE-1 or not containing LINE-1. Microarray expression of mRNAs from the GEO data set (Edgar, 2002; Barrett, 2009) was classified as up-or down-regulated and not up-or not down-regulated depending on the statistical significance determined by the student's t-test. The libraries were GSE6364 (Burney *et al.*, 2007) and GSE4888 (Talbi *et al.*, 2006). These libraries included proliferative (PR), early secretory (ESE), mid secretory (MSE) and late secretory (LSE) phases of endometrium which were processed for total RNA isolation and affymetrix chip hybridization. The student's t-test was performed on all probes. Some probes represented more than one gene (homologous probes). Each gene was counted as up-or down-regulated by expression levels ofunique probes. *P*-values less than 0.05 were observed for represented probes as significantly different secretory phases (ESE, MSE or LSE) and PR.

Identification of up-or down-regulate gene expression analysis

The identification of mRNA of regulated genes from expression microarrays as up- or down-regulated and not up- or not down-regulated were done. The statistical significances were determined by the student's t-test. From microarrays experiment data, each mRNA in the experiments was classified into the following four groups: 1) genes containing LINE-1 and regulated in the array experiments 2) genes containing LINE-1 and not regulated in the array experiments 3) genes without LINE-1 and regulated in the array experiments 4) genes without LINE-1 and not regulated in the array experiments (Figure 1).

The intersection of microarray expression data, which is compatible with Gene Expression Omnibus (GEO), and genes containing the LINE-1 library were done. The number of genes in each subset were compared using the Pearson's chi-square test. The Pearson's chi-square test of independence was determined if two variables (gene groups containing LINE-1 or not) are related (or unrelated) to the interested experiment and provided a 2x2 table for frequency of occurrences. The odds ratio of up- or down-regulated in genes containing LINE-1 was calculated by dividing the probability of the event happening by the probability of the event not happening. The odds ratio of genes containing LINE-1 in up- or down-regulated that is related to the experiment was analyzed. The number of up- or down-regulated genes containing LINE-1 was divided

by the number of not up- or not down-regulated genes containing LINE-1. Then, the number of up- or down-regulated genes that does not contain LINE-1 was divided by the number of not up- or not down-regulated genes that does not contain LINE-1. The odd ratio of more than 1.0 indicated the association between effects of sex steroid hormones and genes containing LINE-1 expression.

Microarray experiment

1) Up – or down regulated genes with LINE - 1	2) Not up – or down regulated genes with LINE - 1
3) Up – or down regulated genes without LINE - 1	4) Not up – or down regulated genes without LINE - 1

Figure 1. The correlation between the expression of genes from microarray libraries and genes containing or without LINE-1. The microarray data were intersected with the genes in four groups as 1) genes containing LINE-1 and regulated in array experiments 2) genes containing LINE-1 and not regulated in the arrays experiment 3) genes without LINE-1 and regulated in the experiment 4) genes without LINE-1 and not regulated in the array experiments. Pearson's chi-square test of independence was determined if two variables (the gene groups of genes containing LINE-1 or not) are related (or unrelated) to the interested experiment and provided a 2x2 table for frequency of occurrences.

RESULTS

Expression of genes containing LINE-1s during the menstrual phase

To investigate sex hormones initiating genes containing LINE-1 expression, we compared gene expression in each phase of the menstrual cycle by using expression arrays from the public database (Gene Expression Omnibus (GEO) with genes containing LINE-1 and genes without LINE-1. Genes containing LINE-1 expression of endometrium in secretory phases (ESE, MSE and LSE) of the menstruation cycle were compared with those in the PR phase. Up- and down-regulate genes were classified by two student's t-tests. Genes containing LINE-1 in each secretory phase were classified as up- or down-regulated when the mean in each phase was statistically higher or lower than the PR phase. From GSE6364, our results showed up-regulated genes with intragenic L1 in the ESE phase were significantly higher than PR phases (p=1.05E-10; odds ratio (OR) = 1.62; 95% confidence interval (CI) =1.40-1.88) while a significance of down-regulated genes was found in MSE phases (OR = 1.15 95% CI = 1.01-1.31, p=2.98E-02) (Figure 2 and Table 1). To validate our results, we performed the similar analysis from another expression microarray of endometrium, GSE4888. The expression of genes containing LINE-1 in three phases of the menstruation cycle (ESE, MSE and LSE phases) were also analyzed and compared to the PR phase. Results were similar to that of up-regulated genes containing LINE-1 which were highly significant in the ESE phase (OR =1.61, 95% CI =1.39-1.87 p=2.22E-10) (Figure 2 and Table 1). Moreover, the results showed that down-regulated genes containing LINE-1 were also found in the MSE phase (OR=1.25, 95% CI =1.10-1.41 p=6.22E-04) (Fig 2 and Table 1)). Although significance of up-regulated genes containing LINE-1 in the MSE phase were found, more samples are needed to clarify. Significance of up-regulated genes with intragenic LINE-1 in the LSE phase was similar to the ESE phase, which presented low levels of E_2 (OR=1.42, 95% CI= 1.19-1.69, p=1.14E-04) (Figure 2 and Table 1).

DISCUSSION

In this study, we investigated for association of hormonal levels in each phase of menstruation and genes containing LINE-1 expression. There are three main interesting observations in the comparison of the PR phase (estrogen dominant) to three phases of SE (progesterone dominant). First, the up-regulation of genes containing LINE-1 had the highest significance in the ESE phase. Second, significance of up and down-regulated genes containing LINE-1 were shown in the MSE phase. Third, up-regulated genes containing LINE-1 expression in the LSE phase was observed.

Generally, FSH and LH from the anterior pituitary need very high level of E_2 in the PR phase to activate ovulation. The increasing amount of E_2 in the PR phase is secreted from the maturing follicle. The appropriate hormonal level is very important to control these functions. Abnormal increase of E_2 levels in the PR phase may cause more proliferation

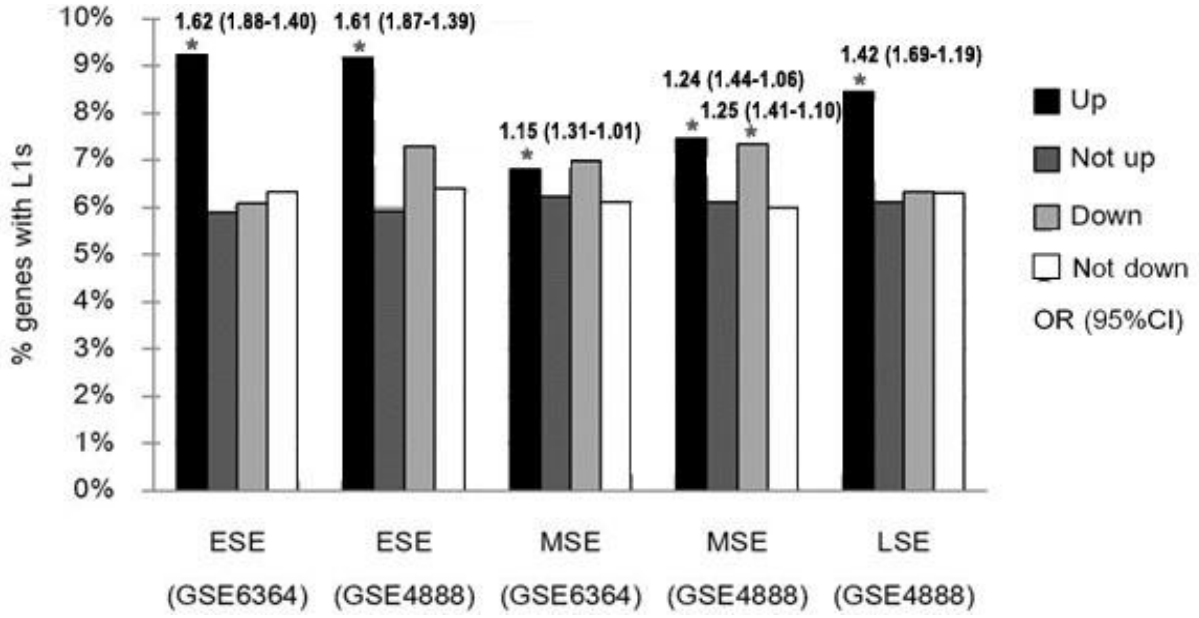

Figure 2 Percentage of up- and down-regulated genes containing intragenic LINE-1 in the early secretory phase (ESE), mid-secretory phase (MSE) and late-secretory phase (LSE) from GSE6364 and GSE4888. GSE6364: the ESE phase: up-and down-regulated genes were presented (*p*-value (odds ratio) =1.05E-10 (1.62) and 6.04E-10 (0.96). The MSE phase: up-and down-regulated genes were presented (*p*-value (odds ratio) =2.23E-01 (1.10) and 2.98E-02 (1.15). GSE4888: the ESE phase: up-and down-regulated genes were presented (*p*-value (odd ratio) =2.22E-10 (1.61) and 2.60E-10 (0.91). The MSE phase: up-and down-regulated genes were presented (*p*-value (odd ratio) =5.66E-03 (1.24) and 6.22E-04 (1.25). The LSE phase: up-and down-regulated genes were presented (*p*-value (odd ratio) =1.14E-04 (1.42) and 9.74E-01 (1.00). Pearson's chi-square test was performed to calculated significant differences ($p<0.05$). Down represents "down-regulated genes", whereas Not down represents "unchanged" or "up-regulated genes". Up represents "up-regulated genes", whereas Not up represents "unchanged" or "down-regulated genes". The asterisk (*) indicates a p-value of <0.05.

of relative cells which originate cancers (Bakkum-Gamez *et al.*, 2011). Then in the secretory phase, high levels of progesterone and lower levels of E_2 (compared to the PR phase) are released from the corpus luteum to produce and maintain endometrium lining. In our results, significance of up-regulated genes containing LINE-1 in the ESE versus PR phase was shown. Levels of E_2 were decreased meanwhile up-regulated genes containing LINE-1 were found. From this result, it seems that the lower level of E_2 in the ESE phase may have had less functions to initiate down-regulated genes, therefore up-regulated genes containing LINE-1 are presented. This may suggest that the high levels of E_2 in the PR phase may function to down-regulate genes containing LINE-1 expression. Moreover, progesterone levels (secreted from the corpus luteum) start to increase in the ESE phase which was the same period that up-regulated genes containing LINE-1 were found. This may indicate that high levels of progesterone may initiate up-regulated genes containing LINE-1. From the analyzed data, the Rapamycin-insensitive companion of Tor (*RICTOR*) was found with intragenic LINE-1 genes and presented up-regulated in the ESE phase. This gene has an associated function related to corpus luteum production. The deletion of *RICTOR* can causes approximately 71% decreasing of the corpus luteum in mice (Chen *et al.*, 2015) which supports our suggestion that high levels of progesterone may initiate genes in the ESE phase.

Second, the significance of up- and down-regulated genes containing LINE-1 were present at the same time that the highest progesterone levels were present. This evidence shows the association of different hormones in the MSE phase to genes containing LINE-1 expression. Down-regulated genes with intragenic LINE-1, suggest that these may be the influence from E_2 levels which still secrete from the corpus luteum and initiate down-regulated genes containing LINE-1. In contrast, up-regulated genes containing LINE-1 in the MSE phase may be initiated by the high level of progesterone. From our results, an example of up-regulated genes containing LINE-1 in the MSE phase is Prolactin receptor (*PRLR*). *PRLR*

can regulate protein-tyrosine phosphatases (PTP) group expression (Edwards *et al.*, 1998). This group of enzymes has an important function related to many pathways, such as signal transduction, cell proliferation and cell cycle control pathways (Shifrin *et al.*, 1997; Stoker, 2005; Östman *et al.*, 2006). Moreover, it was found that increasing PTP activity was high during the secretory phase, but low in the proliferative phase which indicates progesterone-influenced PTP activity (Partanen, 2001). These evidences supported our results that genes containing LINE-1 may be initiated by progesterone in the MSE phase. In the final observation, up-regulated genes containing LINE-1 which presented in the LSE phase, this may indicate that the influence from the decreasing E_2 level causes the degeneration of the corpus luteum.

Table 1 The correlation between menstruation phases and expression of genes containing LINE-1 (GSE6364 and GSE4888) Expression of genes containing LINE-1 in the proliferative phase were compared with those in ESE, MSE and LSE phases. Significant *p*-value and odds ratio were shown in the comparison between up or down-regulated genes containing intragenic LINE-1 in each phase of menstrual cycle. Pearson's chi-square test was performed to calculated significant differences ($p<0.05$). Down represents "down-regulated genes", whereas Not down represents "unchanged" or "up-regulated genes". Up represents "up-regulated genes", whereas Not up represents "unchanged" or "down-regulated genes".

GSE	Menstruation phase	Gene regulation	LINE-1	No LINE-1	Odd ratio	95% CI	p-value
6364	ESE	Up	233	2290	1.62	Lower: 1.40	1.05E-10
		Not up	1067	17023		Upper: 1.88	
		Down	173	2669	0.96	Lower: 0.81	6.04E-01
		Not down	1127	16644		Upper: 1.13	
	MSE	Up	193	2635	1.10	Lower: 0.94	2.23E-01
		Not up	1107	16678		Upper: 1.29	
		Down	340	4540	1.15	Lower: 1.01	2.98E-02
		Not down	960	14773		Upper: 1.31	
4888	ESE	Up	232	2295	1.61	Lower: 1.39	2.22E-10
		Not up	1068	17018		Upper: 1.87	
		Down	179	2281	0.91	Lower: 0.77	2.60E-01
		Not down	1121	16432		Upper: 1.07	
	MSE	Up	220	2732	1.24	Lower: 1.06	5.66E-03
		Not up	1080	16581		Upper: 1.44	
		Down	357	4500	1.25	Lower: 1.10	6.22E-04
		Not down	943	14813		Upper: 1.41	
	LSE	Up	149	1616	1.42	Lower: 1.19	1.14E-04
		Not up	1151	17697		Upper: 1.69	
		Down	282	4182	1.00	Lower: 0.87	9.74E-01
		Not down	1018	15131		Upper: 1.15	

Genes with intragenic LINE-1 present a wide range of biological processes, such as cell differentiation control, cell proliferation, cell homeostasis, genomic stability and especially hormonal response. In normal tissues, LINE-1 elements are less likely to interfere with DNA by acting as hypermethylation (Senthong *et al.*, 2014; Feinberg and Tycko, 2004). Our previous results suggest that LINE-1 methylation levels in normal endometrium did not change even in the different phase of menstruation (about 48%) (Senthong *et al.*, 2014). Therefore, changing of LINE-1 methylation, such as hypomethylation which can decrease host gene expression, was mostly found in cancer (Han *et al.*, 2004; Aporntewan *et al.*, 2011; Xiao-Jie *et al.*, 2016). LINE-1 might be important in controlling gene expression initiated by hormonal levels.

In conclusion, hormonal levels could be an important mechanism to up or down regulate genes containing L1s to control their homeostasis of gene expression. Thus, intragenic LINE-1 may play a role in gene regulation in menstrual cycles.

ACKNOWLEDGEMENTS

Connection up-or down-regulation expression analysis of microarrays extension (CU-DREAM Ex) is a bioinformatics program which is used for the intersection between microarray expression data and genes containing LINE-1 library. The CU-DREAM Ex program and user guideline can be downloaded at

http://pioneer.netserv.chula.ac.th/~achatcha/cu-dream. This research was supported by the National Science and Technology Development Agency (NSTDA) research fund (TG-22-09-55-003D) to SC and the Rachadaphiseksomphot Endowment Fund 2013 of Chulalongkorn University (CU-56-355 HR) to VB.

REFERENCES

Ahmad N, and Kumar R (2011) Steroid hormone receptors in cancer development: a target for cancer therapeutics. Cancer Lett 300: 1–9.

Aporntewan C, Phokaew C, Piriyapongsa J, Ngamphiw C, Ittiwut C, Tongsima S, Mutirangura A (2011) Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. Plos One 6: 1–13.

Bakkum-Gamez JN, Laughlin SK, Jensen JR, Akogyeram CO, Pruthi S (2011) Challenges in the gynecologic care of premenopausal women with breast cancer. Mayo Clin Proc 86: 229–240.

Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Marshall KA, *et al.* (2009) NCBI GEO: archive for high-throughput functional genomic data. Nucleic Acids Res 37: 885–890.

Burney RO, Talbi S, Hamilton AE, Vo KC, Nyegaard M, Nezhat CR, Lessey BA, Giudice LC (2007) Gene expression analysis of endometrium reveals progesterone resistance and candidate susceptibility genes in women with endometriosis. Endocrinology 148: 3814–3826.

Chen Z, Kang X, Wang L, Dong H, Wang C, Xiong Z, Zhao W, Jia C, Lin J, Zhang W, *et al.* (2015) Rictor/mTORC2 in oocytes regulates folliculogenesis and its inactivation causes premature ovarian failure. J Biol Chem 290: 6387–6396 .

Edwards GM, Wilford FH, Liu X, Hennighausen L, Djiane J Streuli CH (1998) Regulation of Mammary Differentiation by Extracellular Matrix Involves Protein-tyrosine. J Biol Chem 273: 9495–9500.

Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207–210.

Feinberg AP, Tycko (2004) The history of cancer epigenetics. Nat Rev Cancer 4: 143–153.

Han JS, Szak ST, Boeke JD (2004) Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature 429: 268–274.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.

Matlik K, Redik K, Speek M (2005) L1 antisense promoter drives tissue- specific transcription of human genes. J Biomed Biotech 2006: 1–16.

Ngamphiw C, Tongsima S, Mutirangura A (2014) Roles of intragenic and intergenic L1s in mouse and human. PloS One 8: 1–8.

Östman A, Hellberg C, Böhmer FD (2006) Protein-tyrosine phosphatases and cancer. Nat Rev Cancer 6: 307-320.

Partanen S (2001) Protein tyrosine phosphatase activity in human endometrium. Reprod Fertil Dev 13: 157–65.

Penzkofer T, Dandekar T, Zemojtel T (2005) L1Base: from functional annotation to prediction of active LINE-1 elements. Nucleic Acids Res 33: 498–500.

Senthong A, Kitkumthorn N, Rattanatanyong P, Khemapech N, Triratanachart S, Mutirangura, A (2014) Differences in LINE-1 methylation between endometriotic ovarian cyst and endometriosis-associated ovarian cancer genes associated with the cis-regulatory functions of intragenic LINE-1 elements. Int J Gynecol Cancer 24: 36–42.

Shifrin VI, Davis RJ, Neel BJ (1997) Phosphorylation of protein-tyrosine phosphatase PTP-1B on identical sites suggests activation of a common signaling pathway during mitosis and stress response in mammalian cells. J Biol Chem 272: 2957–2962

Stoker AW (2005) Protein tyrosine phosphatases and signalling. J Endocrinol 185: 19–33.

Sukapan P, Promnarate P, Avihingsanon Y, Mutirangura A, Hirankarn N (2014) Types of DNA methylation status of the interspersed repetitive sequences for LINE-1, Alu, HERV-E and HERV-K in the neutrophils from systemic lupus erythematosus patients and healthy controls. J Hum Genet 59: 178–188.

Talbi S, Hamilton AE, Vo KC, Tulac S, Overgaard MT, Dosiou C, Le Shay N, Nezhat CN, Kempson R, Lessey BA, *et al.* (2006) Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. Endocrinology 147: 1097–1121.

Wanichnopparat W, Suwanwongse K, Pin-on P, Aporntewan C, Mutirangura A (2013) Genes associated with the cis-regulatory functions of intragenic LINE-1 elements. BMC Genomics 14: 1–9.

Wierman ME (2007) Sex steroid effects at target tissues: mechanisms of action. Adv Physiol Educ 31: 26–33.

Xiao-Jie L, Hui-Ying X, Qi X, Jiang X, Shi-Jie M (2016) LINE-1 in cancer, multifaceted functions and potential clinical implications. J Genet Med 18: 431–439.