The characterization of LINE-1 elements in the human placenta and their role in preeclampsia

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by
Katarina Stevanović, M.Sc
2021

This work was prepared from May 2016 to September 2021 under the supervision of Dr. Zsuzsanna Izsvák (Max-Delbrück Center for Molecular Medicine, Berlin, Germany) in collaboration with Prof. Dr. Ralf Dechend (Experimental and Clinical Research Center, Charité Medical Faculty and Max-Delbrueck Center for Molecular Medicine). All the experiments were conducted in the laboratories of Dr. Izsvák and Prof. Dr. Dechend.

1st Reviewer:

Dr. Zsuzsana Izsvák

Max-Delbrück Center for Molecular Medicine Berlin, Germany

2nd Reviewer:

Prof. Dr. Petra Knaus

Institute for Chemistry and Biochemistry,

Freie Universität, Berlin, Germany

Date of defense: 25.02.2022

I hereby declare that this dissertation was written and prepared by me independently. Furthermore, no sources and aids other than those indicated have been used. Intellectual property of other authors has been marked accordingly. I also declare that I have not applied for an examination procedure at any other institution and that I have not submitted the dissertation in this or any other form to any other faculty as a dissertation.

Table of Contents

1. SUMMARY

Long interspersed nuclear elements (LINE-1), a type of retrotransposon, are genetic elements in the human genome that can self-amplify and integrate back into the genome, increasing their copy number. Their mode of action is called retrotransposition, and it is known to affect the human genome in many ways: generating insertional mutations, genomic instability, alterations in gene expression, and contribute to genetic innovation. These elements are usually silenced in somatic tissue because of their mutagenic potential. I have identified that the evolutionary youngest and human-specific LINE-1 elements are expressed and active in human placental trophoblast cells and that LINE-1 elements are over-expressed and possibly over-active in certain patients of the pregnancy disorder pre-eclampsia (PE).

PE is a pregnancy-specific complex disorder in which a combination of poor placentation and placental stress leads to hypertension and proteinuria in the mother following the 20th week of gestation. PE affects 2-8% of all pregnancies. Despite PE being the leading cause of maternal and neonatal mortality and morbidity, the etiology remains unknown. One major problem in PE research is that PE is a multi origin disorder, exhibiting different phenotypes before the onset of the disease. I hypothesized that LINE-1 overexpression and activity in an immune-mediated subtype of PE patients contributes to PE pathogenesis.

I conducted RNA sequencing in 10 PE and 8 control primary human trophoblast and have identified that human-specific LINE-1 elements, L1HS, are over expressed in 3/10 patients. Using RT-PCR, western blot, and immunohistochemistry methods, I have shown that LINE1s are expressed in primary human trophoblasts at the RNA and protein levels and in all three subtypes of trophoblast cells. Using a LINE-1 retrotransposition reporter assay, I showed that trophoblast cells contain all the necessary conditions to support LINE-1 retrotransposition and that there is greater retrotransposition activity in the PE-model trophoblast cell line. In addition, I identified that LINE-1 over expression in a subset of PE patients induces the expression innate immune response, suggesting a mechanism of early pathogenesis induction in a subtype of PE patients. Lastly, I also showed that in PE, a TE-derived miRNA, has-miR-576-5p, is upregulated, among other newly discovered dyregulated miRNAs. The expression of placenta fuction-relevant genes that are predicted targets of has-miR-576-5p was investigated, but no dysregulation was determined.

Overall, this research revealed novel information about the somatic expression and activity of LINE-1 in the human placenta, and for the first time provided insight into the potential role of LINE-1 in PE pathogenesis.

Keywords: Transposable elements (TE); Long interspersed nuclear elements (LINE-1); microRNA (miRNA); Placenta; Trophoblasts; Pre-eclampsia (PE)

ZUSAMMENFASSUNG

LINE-1-Elemente (Abk. für englisch: long interspersed nuclear elements 1), gehören zu der Klasse der Retroelemente. Sie sind genetische Elemente im menschlichen Genom, die sich (wenn noch aktiv) selbst vermehren. Sie werden in RNA transkribiert, in cDNA umgeschrieben und anschließend wieder in das Genom integriert, wobei sich ihre Kopienzahl erhöht. Ihre Wirkungsweise wird als Retrotransposition bezeichnet, und es ist bekannt, dass sie das menschliche Genom auf vielfältige Weise beeinflussen: Sie führen zu Insertionsmutationen, genomischer Instabilität, Veränderungen der Genexpression und tragen zur genetischen Innovation bei. Aufgrund ihres mutagenen Potenzials wird die Expression dieser Elemente im somatischen Gewebe normalerweise unterdrückt. Ich habe festgestellt, dass die evolutionär jüngsten und humanspezifischen LINE-1-Elemente in den Trophoblastzellen der menschlichen Plazenta exprimiert werden und aktiv sind und dass LINE-1-Elemente bei bestimmten Patienten mit der Schwangerschaftskrankheit Präeklampsie (PE) überexprimiert und überaktiv sind.

PE ist eine schwangerschaftsspezifische komplexe Störung, bei der eine Kombination aus schlechter Plazentation und plazentarem Stress einer übertriebenen mütterlichen Entzündungsreaktion zu Bluthochdruck und Proteinurie bei der Mutter nach der 20. Schwangerschaftswoche führt. PE betrifft 2-8 % aller Schwangerschaften. Die Ätiologie der PE ist nach wie vor unbekannt, obwohl sie die Hauptursache für die mütterliche und neonatale Mortalität und Morbidität ist. Ein großes Problem in der PE-Forschung besteht darin, dass es sich bei PE offenbar um eine Erkrankung mit mehreren Ursachen handelt, die vor dem Ausbruch der Krankheit unterschiedliche Phänotypen aufweist. Ich habe die Hypothese aufgestellt, dass die Überexpression von LINE-1 und seine Aktivität in einem immunvermittelten Subtyp von PE-Patienten zur Pathogenese von PE beitragen.

Ich habe die RNA-Sequenzierung von 10 PE- und 8 Kontroll-Trophoblasten durchgeführt und festgestellt, dass die human-spezifischen LINE-1-Elemente in 3/10 Patienten überexprimiert sind. Mithilfe von RT-PCR, Western Blot und Immunhistochemie habe ich gezeigt, dass LINE-1 in primären menschlichen Trophoblasten auf RNA- und Proteinebene und in allen drei Subtypen von Trophoblastenzellen exprimiert wird. Mithilfe eines LINE-1-Retrotranspositionsreporter-Assays habe ich gezeigt, dass Trophoblastenzellen alle notwendigen Bedingungen für die LINE-1-Retrotransposition aufweisen und dass die Retrotranspositionsaktivität in der Trophoblastenzelllinie des PE-Modells größer ist. Darüber hinaus habe ich die human-spezifische LINE-1 Elemente in SGHPL4-Zellen überexprimiert und festgestellt, dass die Überexpression von LINE-1 in Trophoblasten die Expression von Genen der angeborenen Immunantwortinduziert, was auf einen Mechanismus zur Induktion einer frühen Pathogenese bei einem Subtyp von PE-Patienten hindeutet. Ich habe auch gezeigt, dass in PE eine von TE abgeleitete miRNA, hsa-miR-576-5p, neben anderen neu entdeckten dyregulierten miRNAs hochreguliert ist. Die Expression plazentarelevanter Gene, die mutmaßliche Zielgene von hsa-miR-576-5p sind, wurde untersucht, aber es wurde keine Dysregulation festgestellt.

Insgesamt erbrachte diese Untersuchung neue Informationen über die somatische Expression und Aktivität von LINE-1 in der menschlichen Plazenta und gab erstmals Aufschluss über die potenzielle Rolle von LINE-1 bei der Pathogenese von PE.

2. INTRODUCTION

2.1. Transposable elements (TEs)

Transposable elements (TEs), else known as "jumping genes" or "mobile DNA elements", are repetitive DNA sequences that vary in length, can transpose, or jump, from one location to another within the genome. Due to their repetitive nature, in human genomes TEs were regarded as evolutionary genomic fossils of once active genetic components that have remained in the genomes without serving a particular function and were labeled as "junk DNA"[1, 2]. This belief was re-evaluated when in 1988, the first TE-derived genetic mutation was identified to play a role in human disease [3]. Today it is apparent that TEs make up substantial portions of all eukaryotic genomes in which they exert regulatory power over gene expression by providing gene regulatory sequences such as enhances and promoters. Their perturbation can thus create genetic mutations, and most notably, it has been shown that TEs played a vital role in the evolution of genomic architecture [4].

Barbara McClintock first discovered TEs in the early 1940s by conducting extensive breeding experiments on *Zea Mays* (maize) that produce variable-colored kernels. As illustrated in **Figure 1**, McClintock proposed and demonstrated that upon an ionizing radiation treatment of the kernels, certain genomic elements, which she named Ac/Ds (activator and dissociation) controlling elements, were able to change their locations within and between chromosomes, controlling the gene expression that produces the pigmented phenotype in the kernels [5]. In this work, McClintock has revealed that genomes are not stationary gene catalogs but are dynamic and capable of intra-individual heterogeneity [5]. In the next decade, Britten and Davidson hypothesized that the eukaryotic genomes' repetitive nature could be due to transposition, and they proposed a model where the amplification of repeat elements in the genome could spread "pre-built" regulatory elements to drive the evolution of gene regulatory networks [6-8]. In 1983, Barbara McClintock received the Nobel Prize for physiology or medicine for her discovery of mobile genetic elements. However, it was not until the early 2000s, following the sequencing of the human genome, that we fully understood the vast impact that TEs have had and still exert on genomes.

The completion of the human genome project revealed that TEs make up a large portion of the human genome, about 45%, compared to only about 1% of protein-coding genes, as shown in **Figure 2** [9]. Perhaps that approximation is underestimated since many ancient TEs integrated within the human genome may have accumulated mutations beyond recognition [4]. Novel DNA sequence alignment tools suggest that TEs might make up two-thirds of the human genome [10].

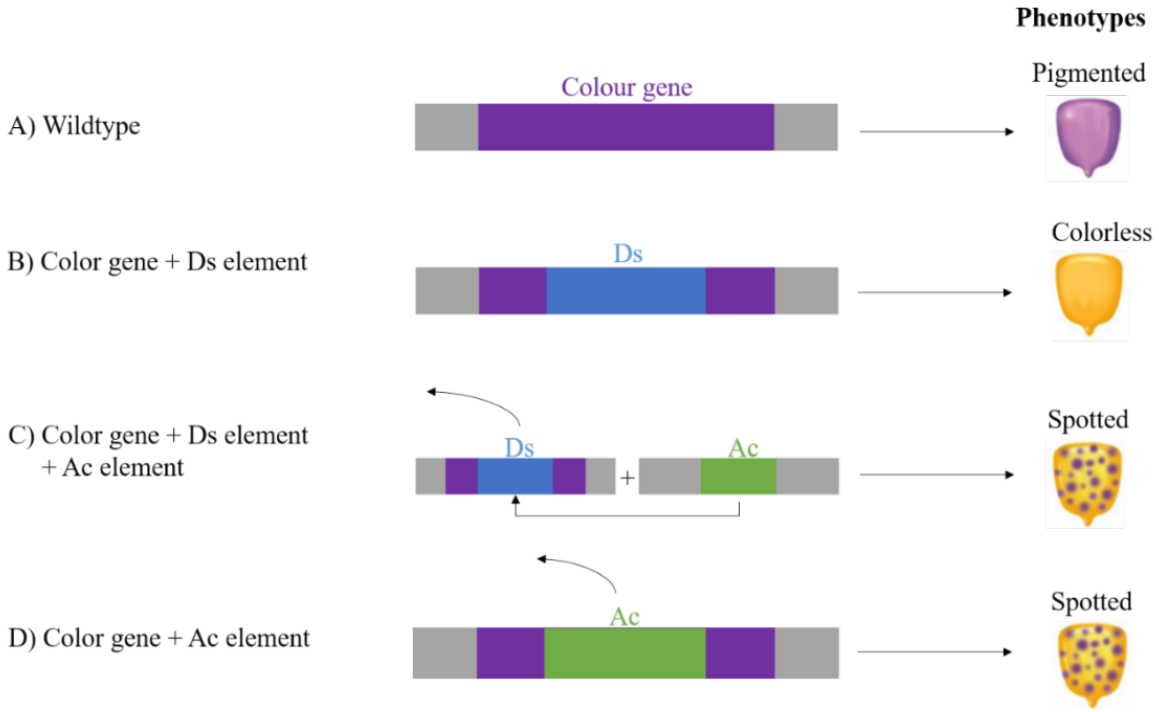


Figure 1: **A summarizing illustration of Barbara McClintock's discovery of transposable elements, termed Ac (Activator) and Ds (Dissociation) elements**. The following observations were made from McClintock's experiments: **A)** The wildtype *color* gene produces a pigmented phenotype of the maize kernel, **B)** the *color* gene is disrupted by an insertion of the Ds (non-autonomous) element within the gene, the functional *color* gene is not expressed and therefore the resulting phenotype is colorless. **C)** the *color* gene is disrupted by insertion of the non-autonomous Ds element within the gene and the autonomous Ac element is in close [9]proximity, then the Ac element can mobilize the Ds element, which can excise from the *color* gene locus; therefore, some cells of the kernel would express the full-length *color* gene resulting in a spotted kernel phenotype. **D)** Similarly, if an autonomous Ac element is present within the *color* gene locus, it can self-mobilize and excise from that position, permitting the spotted kernel phenotype. This figure's illustration was adapted based on figure 14-4 from Introduction to Genetic Analysis, 9th Edition by W.H. Freeman et al., 2007 [11].

2.1.1. TE-driven genome evolution

TEs, as components of genetic information, are usually transferred vertically from parent to offspring via germline cells; therefore, if transposition occurs in the germline, the newly integrated TE copy becomes inheritable [12]. TEs also spread horizontally between species allowing for the colonization of new genomes [12, 13]. Once TEs appeared in genomes, they forced genomes to change in the following evolutionary "arms-race" between TE evolution versus novel host TE repression mechanisms. This antagonistic co-evolution is explained by the Red Queen hypothesis, inspired by Lewis Carroll's - Through the looking glass, the Red Queen's statement to Alice that it "takes all the running you can do to keep in the same place." In other words, continuous evolution is needed to maintain its fitness relative to the systems that it is co-evolving with [14]. Since uncontrolled TE expression and transposition could ultimately be detrimental to the host, host genomes have developed mechanisms to repress the expression of TEs, at the transcriptional, post-transcriptional, and translational levels [15] [16]. However, to self-propagate, TEs have also mutated throughout evolution to evade the host's newly evolved TE repression mechanisms, in turn pressuring the host further to evolve the TE repression mechanisms [12]. TEs are not segregated into a separate compartment of the genome but am scattered within and around most genes [17]. Consequently, TEs transformed genomes of their hosts by causing genome instability (genomic expansion and deletion), mutation, and novel gene regulation mechanisms, as shown in **Figure 3** [12, 18-20].

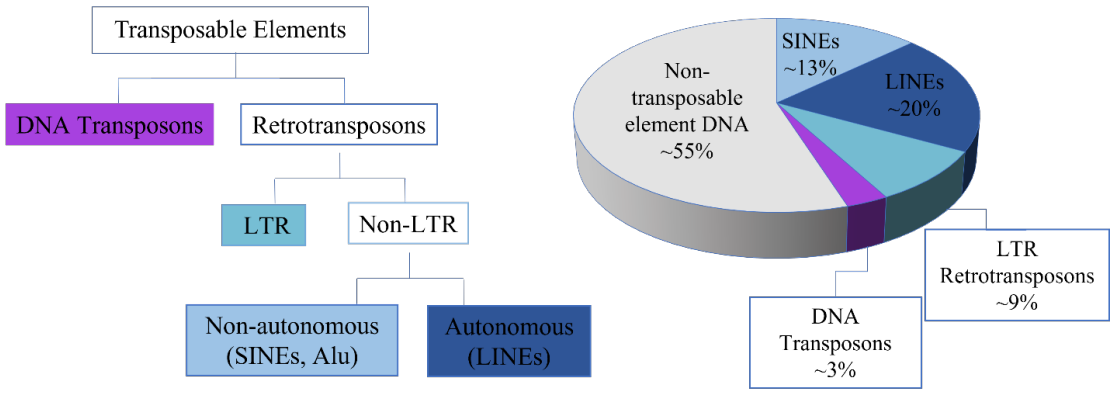


Figure 2: An overview of different types of TEs, their modes of transposition, and their contribution to the human genome. A) A tree diagram representing the different classes and subclasses of TEs. A pie chart of the human genome composition, emphasizing the percentage of

different sub-classes of transposable elements. Retrotransposons are represented by various blue shades (SINEs, LINEs, and LTR retrotransposons), and DNA transposons are represented with purple. The chart data was obtained from Lander et al. 2001 [9].

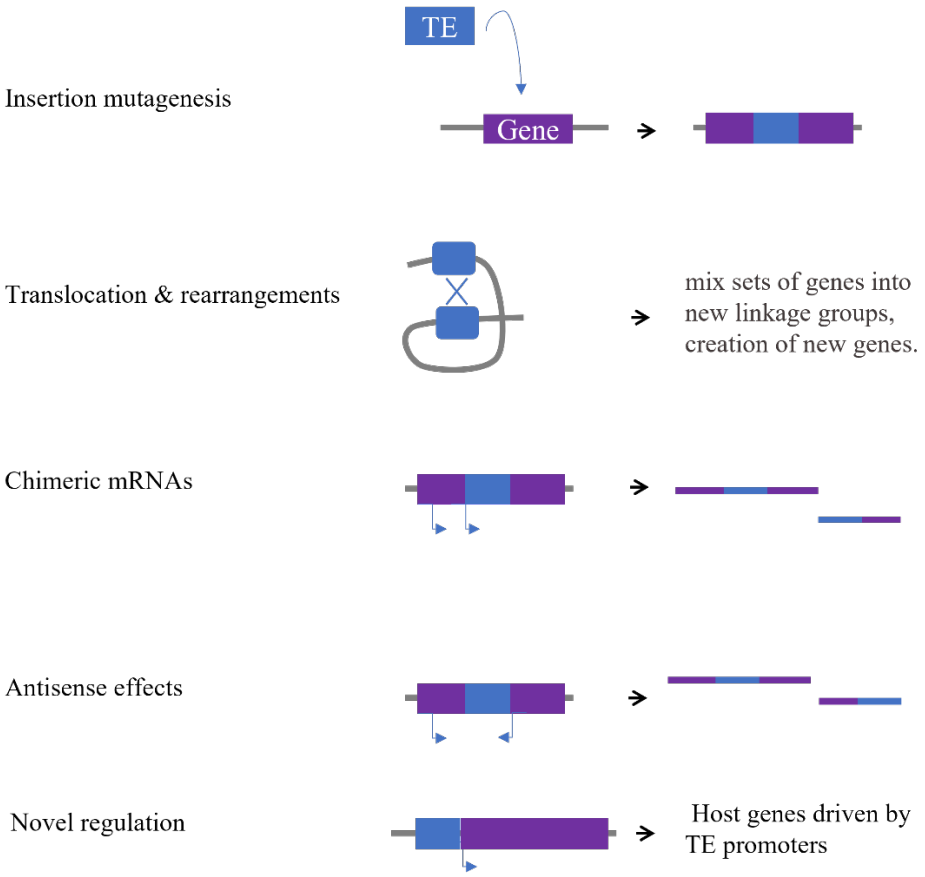


Figure 3: An illustration of the possible TE effect on the structure and expression of the genome. This figure was adapted from Yoder et al. 1997 [17].

2.1.2. Different types of TEs

There are two major classes of TEs, retrotransposons and DNA transposons, distinguished by their unique mobilization mode [19]. Retrotransposons are DNA sequences that transpose via an RNA intermediate, using a "copy-and-paste" mechanism. As illustrated in **Figure 5**, they reverse-transcribe their mRNA products into cDNA and insert this new copy into the genome, increasing in copy number [21, 22]. DNA transposons move by a "cut-and-paste" mechanism, except for

helitrons which use the rolling circle replication [23]. They excise themselves and re-insert into a new genomic site, increasing their copy number by homologous recombination [22]. DNA transposons were formerly active during primate evolution. However, due to an accumulation of many mutations within the DNA transposons and due to a lack of polymorphisms in the human populations, it is suggested that DNA transposons have not been active in the human lineage for over 40 Myr [9, 24].

2.1.2.1. Retrotransposons

Retrotransposons are the major class of TEs in the human genome [9]. As shown in **Figure 2**, retrotransposons are further divided into two subclasses: sequences containing long terminal repeats (LTRs) and non-LTR-containing sequences.

The LTR retrotransposons, or endogenous retroviruses, are derived from exogenous retroviruses that have infected the germline several times during primate evolution [25]. Original LTR retrotransposons were retroviral-like in structure. Those LTRs are eukaryotic-like promoter sequences and contain an RNA polymerase II binding site, as well as cis-regulatory sequences with which viral and cellular regulatory factors can interact [26]. They also encode for the three genes that are needed for viral replication: group-specific antigen (*gag*), polymerase (*pol*) - which also contains a reverse transcriptase and an endonuclease, and an envelope domain (*env*). However, due to mutations, they have lost their ability to replicate and infect [24]. Nevertheless, LTR sequences can affect host gene transcription through enhancers and promoters that have contributed to the spatiotemporal expression of their neighboring genes [27].

The second type of retrotransposons, the non-LTR retrotransposons, are further distinguished into autonomous and non-autonomous groups. Autonomous elements such as the LINE1s (Long Interspersed Nuclear Elements) produce their proteins for mobilization, and non-autonomous elements such as SINEs (Short Interspersed Nuclear Elements) mobilize in-trans by the expression of mobilizing proteins from the autonomous elements [28]. The non-LTR retrotransposons are further subclassified based on their evolutionary age. In the human genome, there are still structurally intact LINE-1, SINE, Alu, and SVA elements capable of mobilization [29-33].

2.1.2.2. LINE-1 elements

Based on the conserved sequences of the reverse transcriptase-bearing group II introns of bacteria and mitochondria, LINE-1 elements are suggested to be ancient elements that have been present in eukaryotic genomes throughout evolution [34]. Since the emergence of the apes in the past 25 million years, LINE-1 has amplified and evolved, generating five distinct LINE-1 subfamilies, L1PA5 to L1PA1 [35-37]. L1PA1 is generally referred to as L1Hs, the human-specific LINE-1 subfamily [35]. Due to this massive mammalian expansion, about 500,000 copies of LINE-1s are found in the human genome, composing approximately 17% of the genome [4, 9]. Most of these LINE-1 elements have accumulated mutations (such as frameshifts, 5' truncations, inversions, substitutions, INDELs, alterations in either of the ORFs) which render the elements incompetent [38-40]. However, there are about 80-100 copies in the human genome of the evolutionary youngest LINE-1 subfamily, the human-specific L1Hs, that are full-length and retrotranspositionally active [9, 37, 41, 42]. However, further studies have revealed that most LINE-1s have low retrotransposition efficiency; therefore, only about 6-10 copies of the above-mentioned LINE-1s, named "hot L1s", conduct the bulk of LINE-1 retrotransposition in the human genome [41, 42]. These active elements can be distinguished based on specific diagnostic nucleotides in the LINE-1 sequence, such as "ACA" at positions 5930-5932 and a "G" at 6015 [30]. In contrast, older and inactive elements have "GAT" and an "A" at the respective positions [30]. L1Hs elements are differentially present in individual genomes but absent from the Human Genome Reference sequence; therefore, it is possible to segregate populations based on L1Hs insertion polymorphisms [42-47]. This indicates that LINE-1 elements are important in generating inter-individual genetic variation, cancer, and other genetic diseases.

2.1.2.3. LINE-1 structure

Full-length LINE-1 elements are approximately 6000 base-pairs long that contain a sense RNA polymerase II promoter and an antisense promoter in the 5'UTR including a usually methylated CpG island, three non-overlapping open reading frames (ORF1, ORF2, and ORF0), a 205 bp 3' UTR, and a polyA tail of a variable length, as shown in **Figure 4** [47-52]. The sense promoter generates the LINE-1 ORF1 and ORF2 mRNA transcripts. The ORF1p encodes for an

approximately 40 kDa protein that contains an RNA recognition domain and functions as a nucleic acid chaperone [53-55]. ORF2p encodes for a multifunctional 150kDa protein that contains an endonuclease (EN), a reverse transcriptase (RT), and C-terminal cysteine-rich domains [56-58]. The LINE-1 antisense promoter can generate L1-gene chimeric transcripts that include neighboring exon sequences [59]. It also drives the expression of the recently discovered primate-specific ORF0 [52, 59, 60]. There are only about 781 loci in the human genome that could encode for ORF0 [52]. This gene product is not clearly understood yet. However, it is known that ORF0 localizes near nuclear bodies and enhances LINE-1 retrotransposition *in vitro,* but its mechanism of action is not yet clearly understood [52].

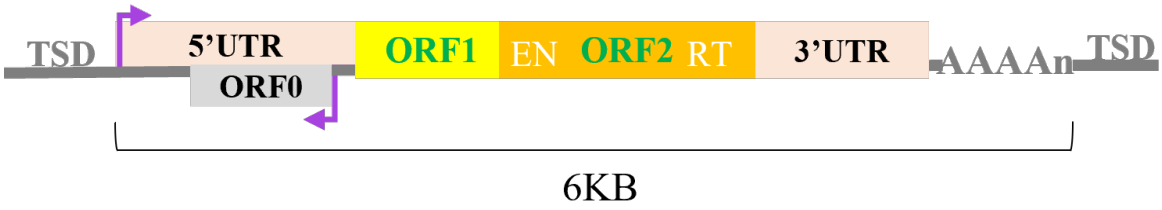


Figure 4: An illustration of the features of a LINE-1 locus. Legend: purple arrows = the sense and non-sense promoters, peach rectagles = untranslated regions, yellow rectangle = open reading frame 1, orange rectangle = open reading frame 2 (including the two domains, endonuclease and reverse transcriptase), AAAAn = polyadenylation tail of a undefined length, TSD = target-site duplication regions.

2.1.2.4. LINE-1 retrotransposition

LINE-1 retrotransposition begins with the transcription of an existing genomic full-length LINE-1, into a bicistronic mRNA by RNA polymerase II [49, 61]. The LINE-1 mRNA is then translocated from the nucleus to the cytoplasm, where it is translated into the two proteins, ORF1p and ORF2p, that are both required for LINE1 activity [51]. The LINE-1-derived proteins have a *cis* preference, meaning that they preferentially associate with their mRNA by binding to the stem-loop structure of the 3' tail, leading to a formation of a stable ribonucleoprotein particle (RNP) which is also necessary for retrotransposition [62-64]. The exact composition of the RNP has not been determined, but it is known that it contains at least one LINE-1-derived mRNA, a homotrimeric ORF1p, and an ORF2p [55, 65]. While ORF1p proteins are abundantly available, ORF2p is believed to be translated by an unconventional mechanism at about one molecule per

RNA [52, 66]. Next, the LINE1-RNP complex enters the nucleus by an unknown mechanism where LINE-1 retrotransposition events accumulate in the AT-rich regions of the human genome [67]. The insertion of a new LINE-1 copy into the host genome occurs via the Target-Primed Reverse Transcription (TPRT) mechanism [68, 69]. TPRT is initiated by the EN domain of ORF2p, making a nick on one strand of the double-stranded genomic DNA between the sequences TTTT/AA (referred to as the EN recognition sequence) and exposing the thymine 3'OH group. This provides a primer for the RT to initiate reverse transcription of the LINE-1 mRNA (the template) into cDNA starting from the polyA tail [56]. To synthesize the LINE-1 cDNA for the second strand of the double-stranded genomic DNA, the second strand is nicked, however usually downstream from the cleavage of the first strand, and this distance results in target site duplications (TSDs) of approximately 15 bps in length [70]. Often, the internal sequences of the newly synthesized cDNA for the first strand of the genomic DNA anneal with the 5' end of the EN recognition sequence and terminate the reverse transcription, which possibly enhances the synthesis of the second strand cDNA [70]. This results in TPRT decay in which most new LINE-1 copies are 5' truncated and are therefore inactive for future retrotransposition [39, 71].

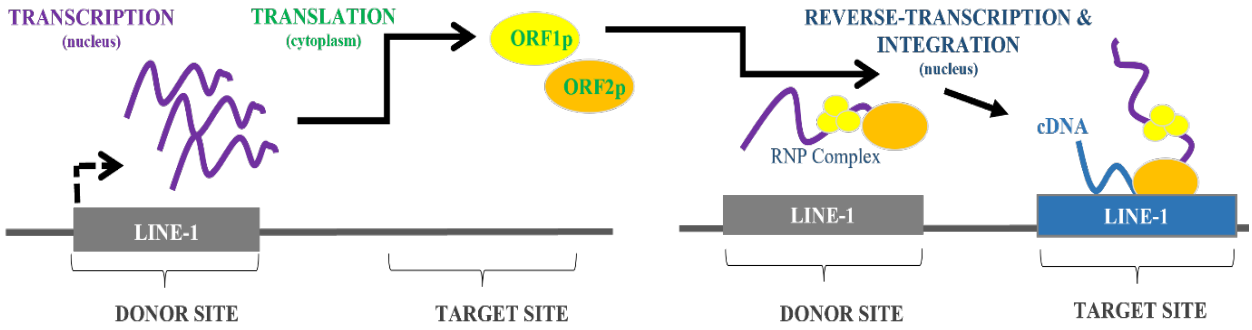


Figure 5: **An illustration of the LINE-1 retrotransposition mechanism.** The copy-and-pasted mechanism: the integrated retrotransposon in the genome is transcribed into mRNA. In LINE1, the bicistronic mRNA is then translated into two proteins, ORF1p and ORF2p. ORF1 contains an RNA recognition domain and binds to a newly transcribes retrotransposon-derived RNA molecule. The ORF2p protein contains an RNA reverse-transcriptase and nuclease activity domain, and there it reverse-transcribes the retrotransposon-derived RNA molecule to cDNA.

2.1.2.5. LINE-1 regulation

To limit the activity of potentially mutagenic LINE-1s, cells have acquired several defense mechanisms interrupting retrotransposition at the levels of transcription, post-transcription, and post-translation. The main method of studying LINE-1 regulation in humans is by using LINE-1 reporter vectors in cell lines; however, most of these studies have been done *in vivo* in mice.

The host cells need to have a suitable environment for LINE-1 expression. For transcription initiation, several important transcription factor binding sites have been identified within the LINE-1 5' UTR, such as YY1, the SRY family, RUNX3, and p53 [72-75]. The expression of these transcription factors is either required or essential for regulating LINE-1 transcription in the different human cell types [76]. Depending on the tissue type, other transcription factors may be involved in regulating LINE-1 transcription [77]. However, the major mechanism for LINE-1 regulation is transcriptional silencing by overlapping epigenetic modifications including DNA methylation, histone deacetylation, and histone methylation [17, 78-80].

The LINE-1 5'UTR contains a CpG island which is most often strictly methylated to suppress LINE-1 expression [81]. As LINE-1s are so widespread in the human genome and heavily methylated in somatic cells, the methylation status of the LINE-1 5'UTR is commonly used as a readout of the methylome status in research. The main mechanism of DNA methylation, and therefore LINE-1 repression, is by DNA methyltransferases (DNMTs). It has been shown that mutant mouse embryos for the maintenance methylation enzyme *DNMT1*, and the *de novo DNMT3a* and *DNMT3b* double-mutant mice exhibited decreased levels of LINE-1 methylation [82, 83]. Also, the deletion of DMNT3L in mice, a protein required for *de novo* methylation without methyltransferase activity, resulted in a global loss of *de novo* methylation of LINE-1 [84]. Interestingly, Drosophila, which suffer from abundant TE-mediated mutations, do not have DNA methylation[17]. In fact as TEs are rich in CpG dinucleotides and heavily methylated, the majority of methylated cytosines in the genome actually are within TEs [17]. It has also been demonstrated that the methyl CpG binding protein 2, MeCP2, binds to the methylated CpGs of the LINE-1 promoter to recruit histone deacetylases, contributing to an inactive chromatin structure [85]. Depletion of H4K20me3 is closely associated with DNA hypomethylation in LINE-1, resulting in genomic instability[86]. Ren et al. have demonstrated a direct link between

H4K20me3 and DNMT1-mediated DNA methylation, providing mechanistic insights into DNA methylation maintenance of LINE-1 [87].

Moreover, a vast repertoire of post-transcriptional defenses against LINE-1s exists. This includes the apolipoprotein B mRNA editing enzyme (APOBEC) proteins, a vertebrate-specific family of cytidine deaminases. APOBEC enzymes edit mRNA sequenced by deaminating cytosine to uracil, which produces a stop codon and truncated protein [88]. There are 7 APOBEC proteins (A3A-D, A3F, A3G, and A3H), and most are known to inhibit LINE-1 retrotransposition to varying degrees, with A3A being the most effective [77, 89]. Another mechanism is RNA interference (RNAi) -mediated LINE-1 silencing. PIWI-RNAs (piRNAs) specifically silence LINE-1s in the germline [77]. It has also been demonstrated that microRNAs (miRNAs), particularly mir-128 and let-7, can silence LINE-1s in HeLa cells and induced pluripotent stem cells (iPSCs) [77, 90] [16]. Knockdown experiments of DICER1 or AGO2, both involved in the miRNA processing pathway, show an increased rate of LINE-1 retrotransposition *in vitro* [91].

Little is known about the post-translational regulation of retrotransposons. The only reports are regarding the phosphorylation of LINE-1 ORF1p at critical phospho-acceptor residues (two serines and two threonines) by the proline-directed protein kinase (PDPK), which is required for L1 retrotransposition [92].

As previously stated, there is an ongoing evolutionary arms-race between TE evolution and host TE repression mechanisms. There are existing repression mechanisms in humans that have specialized to regulate sub-families of LINE-1s that have evolved. The middle-aged LINE-1 subfamilies, L1PA4 and L1PA5, are suppressed explicitly by the KRAB-associated protein 1 (KAP1), also known as TRIM28 [93]. In this mechanism, KRAB-zinc finger proteins (KRAB-ZFPs) transcriptionally silence target sequences via the zinc-finger binding domains, which then bring in proximity the repressive Kruppel-associated box (KRAB) domain [94]. KRAB-ZFPs also interacts with KAP1, which recruits the H3K9me3 methyltransferase, SETDB1, to silence the targeted DNA sequence [94]. In embryonic stem cells (ESCs), the depletion of KAP1 upregulated LINE-1 expression [93]. The specificity of KAP1 for the older LINE-1 elements is due to evolutionary mutations in the younger LINE-1s within the zinc-finger recognition sequences. In contrast, young and full-length LINE-1s have been demonstrated to be repressed explicitly by MORC2, which belongs to the human silencing hub (HUSH) complex that is further composed of

the proteins TASOR, MPP8, and PPHLN1, and SETDB1 [95-97]. MORC2 binds selectively to the 5'UTR of full-length young LINE-1s. Further, the HUSH complex can recruit the protein SETDB1 to the heterochromatic sites to add additional H3K9me3, therefore spreading heterochromatinization [98]. The HUSH complex's interaction with MORC2 is required for its ATPase activity to facilitate the gene silencing [99]. As this mechanism is proposed to target the youngest and therefore potentially active LINE-1 elements, it is the most relevant mechanism to focus on for studying the role of LINE-1 elements in disease and development.

2.1.2.6. The role of LINE-1 in human disease and development

LINE-1s have been demonstrated to play a role in various diseases and are now indicated to contribute to normal development and aging. The first report of a mutagenic LINE-1 insertion was in a patient with hemophilia A, in which a new exonic LINE-1 insertion had occurred in the X-linked blood clotting factor VIII gene [3]. Since then, more than 100 cases have been identified in human genetic disorders caused by *de novo* LINE-1 insertions, including β-thalassemia, Duchenne muscular dystrophy, retinitis pigmentosa, and cancer [100-103]. In one colorectal cancer case, a LINE-1 insertion disrupted the tumor-suppressor adenomatous polyposis coli (*APC*) gene in the tumor cells; however, it was absent in the surrounding tissue of the patient, indicating that LINE-1 can mobilize in some somatic cells [104]. Moreover, and independent of retrotransposition, when LINE-1s are too highly expressed, both transcripts and their truncated proteins can cause varying levels of cell toxicity and DNA damage [105]. Also, overexpression of the L1-ORF2p can cause double-stranded breaks to induce cell senescence and apoptosis [106]. Finally, a loss of LINE-1 CpG-island methylation, which may facilitate genetic instability by DNA breaks and chromosome translocations, is a common feature of cancerous and aging cells [107, 108]. Since LINE-1 elements are often found near genes, a change in their chromatin status, such as LINE-1 mediated heterochromatinization due to a novel LINE-1 insertion, could repress any surrounding genes as well [4, 109]. In cancer cell lines, it has been shown that inactivation of epigenetic silencing leads to the reactivation of LINE-1s, leading to the activation of the innate immune system, and a worsening cancer progression [110-112].

Similarly, in germ cells and early embryogenesis, which have less epigenetic repression, LINE-1s are active during normal development [113, 114]. The inheritance of new germline LINE-

1 integrations is responsible for the continuing increase of the number of LINE-1 insertions in the human genome and allows for the propagation of the LINE-1 elements. Up to 5% of newborn children have a new retrotransposon insertion during early embryogenesis, contributing to inter and intra-individual heterogeneity [114]. Somatic LINE-1 activity was also demonstrated to occur in human fetal neural progenitor cells and the adult brain, suggesting that most human brain neurons may contain a unique genome [115-117]. The benefit and functionality of somatic mosaicism remains unclear, but it is suggested to add to an organism's genetic plasticity and overall fitness. However, an excess of LINE-1 retrotransposition is detrimental. For example, in the brain, it is speculated that increased LINE-1 insertions could contribute to neuronal decline since de-repression of LINE-1 occurs in a variety of neurodegenerative diseases, such as Rett Syndrome (associated with mutations in the LINE-1 regulator MeCP2), Ataxia telangiectasia disorder, frontotemporal lobar degeneration, and schizophrenia [118-122]. A few other neuropathological states, such as in addictive drug use and post-traumatic stress disorder, have been shown to involve dysregulated LINE-1 expression due to epigenetic changes in the associated brain cells [123, 124].

Overall, LINE-1 elements have been shown to be active during pre-implantation embryos, ESCs, iPSCs, neuro progenitor cells, brain cells, and cancerous cells [125]. Conversely, LINE-1 elements are generally highly suppressed in somatic tissues.

However, one of the somatic tissues showing high levels of LINE-1 expression, and the subject of this thesis, is the human placenta. Firstly, the human placenta undergoes global hypomethylation in the first trimester of pregnancy, providing a suitable environment for expressing LINE-1 elements [126]. Using RNA-sequencing, my colleagues, Zadora et al., have shown that the L1HSs are highly expressed in isolated cytotrophoblasts of primary human placentas and are over-expressed in one pre-eclamptic patient, suggesting that LINE-1s may play a role in pre-eclampsia (PE) pathogenesis [127].

2.1.3. The role of TEs in the origin and evolution of miRNAs

Micro RNAs (miRNAs) are short, non-coding small RNA molecules that are approximately 22 nucleotides long and are responsible for post-transcriptional regulation of about 30% of all protein-coding genes in mammals [128, 129]. The biogenesis of miRNAs involves the transcription of pri-miRNAs from the genome by RNA polymerase II [129, 130]. The pri-miRNA

is processed by the enzymes Drosha and Dgcr8 into pre-miRNA in a stem-loop structure, which is then exported to the cytoplasm and finally processed into a mature miRNA by the enzyme Dicer yielding a short double-stranded RNA [129, 130]. The single-stranded miRNA associates with Argonaute (Ago) proteins that provide stability for the miRNA. Together, the miRNA and Ago proteins form the RNA-induced silencing complex (RISC) complex, responsible for mRNA decay and translation repression [129-131]. The miRNAs target specific genes by binding, often partially, to a complementary seed sequence on the mRNA usually found on the 3'UTR and is about 5-7 nucleotides long. Most often, miRNAs function through their combined action, they often have a large number of targets, and commonly they amplify their effect by targeting multiple genes in a common signaling pathway [130, 132].

TEs are hypothesized to provide a natural mechanism for the formation of new miRNAs, for example, by converging TEs and by TE mutagenesis [133-139]. **Figure 6** illustrates a possible cellular event responsible for TE-derived miRNAs and their target genes [139]. In line with this, Qin et al. compared the locations of TEs with those of pre-miRNAs on query sequences, those that matched were termed MDTE, miRNAs derived from TEs [138]. They found that 19.84% of all miRNAs overlap with TEs in humans, and they show unique relationships to their related TEs [138]. Wei et al. have developed an MDTE database (MDTE DB), consisting of 2583 miRNAs derived from TEs that allows for storage, search, and analysis of MDTEs [137].

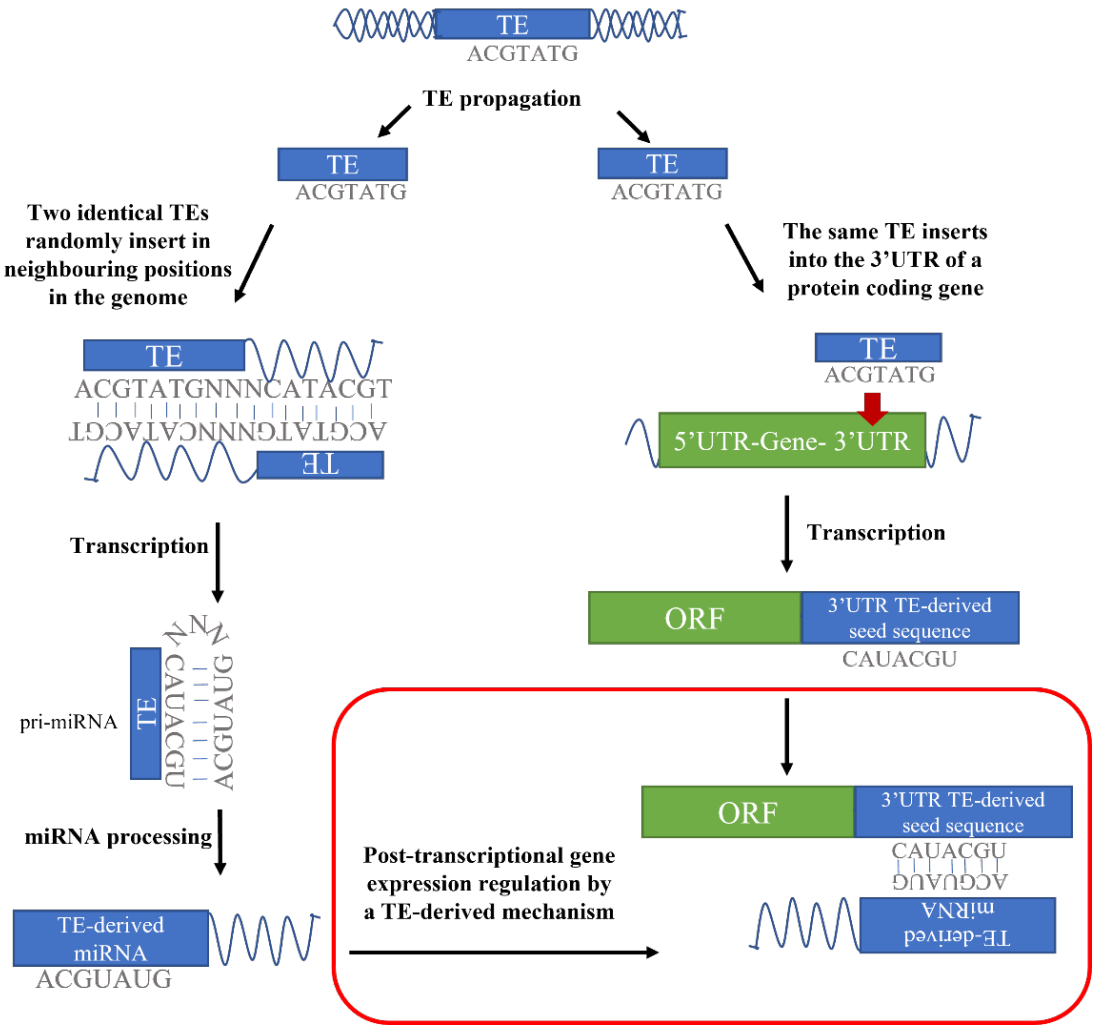


Figure 6: An illustration of a potential mechanism for TE-derived miRNA-mediated gene expression regulation. An active TE, containing the sequence "ACGTATG" transposes and propagates across the genome, randomly integrating. If this TE integrates into close vicinity to itself in the genome, that event can lead to the formation of a miRNA (left). Similarly, and through time, this same TE can insert itself into a non-coding sequence of a gene elsewhere in the genome, providing a seed sequence for the newly created miRNA that is derived from the same TE (right). Outlined in red is the combination of the two series of events occurring that lead to the formation of a miRNA gene regulatory network that is derived from a TE. This figure was adapted from Roberts et al. 2014 [139].

2.1.3.1. miRNAs involvement in human disease

Altered miRNA expression is associated with several pathologies. Particularly in cancer, dysregulated miRNAs play causal roles in specific malignancies [140]. Many miRNAs were found to play critical roles in key biological processes such as cell division, metabolism, immunity, angiogenesis, and mobility, indicating that abnormal miRNA expression can significantly affect those critical processes, resulting in various pathological outcomes [141].

Over 600 miRNAs are predominantly or exclusively expressed in the placenta and play a developmental role throughout gestation [142]. They are mostly encoded in clusters in the genome and act synergistically [143]. The three largest pregnancy-related miRNA clusters are the chromosome 19 miRNA cluster (C19MC), the chromosome 14 miRNA cluster (C14MC), and the miR371-3 cluster that is also found on chromosome 19[143]. In the above-mentioned pregnancy disorder, PE, miRNAs are increasingly being reported as contributors to the disease and are of great interest for research due to their potential as non-invasive biomarkers of PE since they can be detected in the maternal plasma and sera [143-146]. Among many others, the most well studied and most highly expressed PE-related miRNA is hsa-miR-210, which targets HIF-1α, potentially leading to the hallmark hypoxic conditions of PE. More complexly, has-miR-210 was upregulated in severe PE but downregulated in mild PE [147]. Several individual miRNA-gene targets have already been revealed for PE; however, the search for an early detection biomarker remains.

2.2. Human pregnancy

Pregnancy, also referred to as gestation is defined as when a fetus develops in the womb[148]. During a woman's menstrual cycle, ovarian sex hormones, mainly progesterone, and estrogen, mediate the body's preparation for pregnancy. Around the fourteenth day of the menstrual cycle, at the peak of estrogen expression, an ovum (a mature oocyte) is released from an ovary, which is followed by a rise in progesterone that thickens the endometrium (decidualization) for conceptus receptivity [149, 150]. A released ovum (a female-origin haploid gamete) travels from the fallopian tube towards the uterus, where if a spermatozoon (male-origin haploid gamete) is present, it will fertilize the ovum to create a zygote, a single diploid cell containing the maternal and paternal DNA [151]. While traveling towards the decidualized endometrium (the decidua) in the uterus , the pre-implantation conceptus divides from the single totipotent cell to the morula until the blastocyst stage, as illustrated in **Figure 7** [151]. The appearance of a fluid-filled inner cavity marks the transition from the morula to the blastocyst, accompanied by cellular differentiation [152]. The blastocyst is composed of two major cell types, the non-polarized inner cell mass (pluripotent stem cells that give rise to all three germ layers of the embryo) and the enveloping polarized trophectoderm (TrE) cells. TrE cells that give rise to the extraembryonic tissue such as the different sub-populations of trophoblast cells and the umbilical cord [153]. After the sixth day of fertilization, the trophectoderm of the blastocyst lays adjacent to the decidua, to which it adheres and stably anchors. This is the location where the development of the fetus and the placenta continues until birth [152, 154]. The average length of a human pregnancy, or gestational age (GA), is about 280 days or 40 weeks, counting from the last menses, subdivided into three trimesters based on distinct developmental stages [148]. During pregnancy, the woman's body undergoes substantial physiological changes that affect every organ in the body. These changes include hematological changes, increased cardiac output, renal anatomy and function, the endocrine system, altered metabolism, and bone density changes [155]. In most uncomplicated pregnancies, these changes resolve after delivery [155].

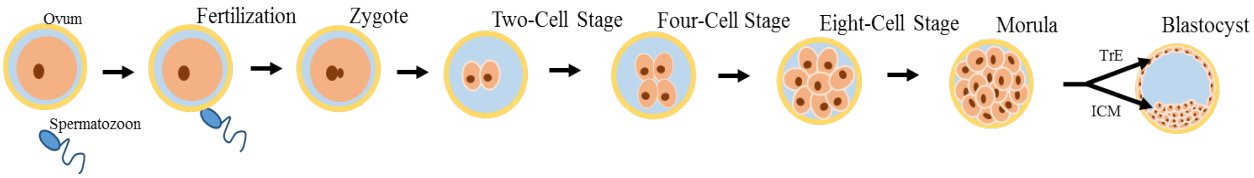


Figure 7: Fertilization and Pre-implantation Conceptus Development. Upon fertilization of an ovum by a spermatozoon, the single-celled zygote continuously divides by mitosis to form the morula (a mass of 12-16 cells), where cells begin to differentiate into the trophectoderm (TrE) and the inner cell mass (ICM). The polarization of the cells leads to the formation of the blastocyst.

2.2.1. The human placenta

The human placenta is a transient hemochorial organ that exists only during pregnancy, influencing pregnancy progression and the lifelong health of the mother and the baby [156]. The placenta embeds in the maternal decidua during early pregnancy, and the placental extra-villous trophoblast cells invade the maternal spiral arteries, creating a maternal-fetal interface connected to the fetus by the umbilical cord and to the mother by direct contact with the decidua and the uterine spiral arteries. This establishes two circulation systems: the fetoplacental blood circulation and the uteroplacental blood circulation [151, 157]. The feature that makes human placentation unique is the highest degree of trophoblast invasion into the maternal decidua versus the invasion that takes place in other placenta-bearing species [158], even though extensive trophoblast invasion has also been reported in chimpanzees and gorillas [159].

During pregnancy, the placenta grows until a term weight of nearly 500g; it is about 15–22cm long in diameter, about 2 cm thick, and has a surface area of approximately 15 m^2 [160, 161]. The primary function of the placenta is to produce and secrete endocrine hormones and growth factors that support and promote pregnancy and fetal growth. To accomplish that, it has exchange to oxygen and carbon dioxide between the mother and the baby, absorb nutrients for the maternal circulation that nourish the developing fetus, excrete fetal waste, protect the semi-allogeneic fetus from the maternal immune system, create a protected microenvironment from external stresses and pollutants, and to initiate parturition [162-168].

2.2.1.1. Human placenta development and structure

Following the blastocyst formation of early embryo development, further differentiation of the TrE cells give rise to cytotrophoblasts (CTBs), also referred to as trophoblast stem cells [169]. These cells express CTB marker genes such as *TFAP2C*, *GATA3*, and *KRT7[169]*. CTBs give rise two types of trophoblast subpopulations that are morphologically and functionally distinct: (i) syncytiotrophoblasts (STBs) and (ii) extra-villous trophoblasts (EVTs) [169]. The different trophoblast cell types are the main constituents of the placenta, and their coordinated proliferation and differentiation create the chronic villi structure of the placenta. STBs are multinucleated cells generated by a high degree of differentiation and syncytial fusion of villous cytotrophoblasts [170]. STBs form the exterior cell layer of the placenta and produce important endocrine factors, such as hormones and growth factors, for the development of the placenta and the fetus throughout pregnancy [170].

Chorionic villi development occurs in three stages, as shown in **Figure 8**. Firstly, CTBs proliferate and differentiate to form primary villi consisting of an overlaying STB layer and underlying villous CTBs. Primary villi are finger-like outgrowths of only trophoblast cells proliferating laterally into the intervillous space to develop the villous trees [171]. Secondary villi are formed by CTB proliferation and villi branching, and a primary mesoderm grows within each villus; therefore, the villi do not contain only trophoblast cells any longer [171]. Lastly, with the formation of first fetoplacental blood vessels, two types of tertiary chorionic villi form. Firstly the anchoring villi that attach to the decidua, and secondly the floating villi that float in the intervillous space [171].

The CTBs within the anchoring villi break through the syncytium to form trophoblastic cell columns, which differentiate into EVTs with an invasive phenotype [172]. They are sub-divided into an interstitial (iEVTs) and an endovascular (eEVTs) type. The iEVTs migrate and invade the uterine interstitium and communicate with cells found in the maternal environment, such as uterine natural killer cells (uNK), macrophages, and decidual stromal cells [170]. The eEVTs colonize and remodel maternal spiral arteries, transforming from high-resistant, low-flow maternal vessels into large, dilated, low-resistant, and high-flow ones [170].

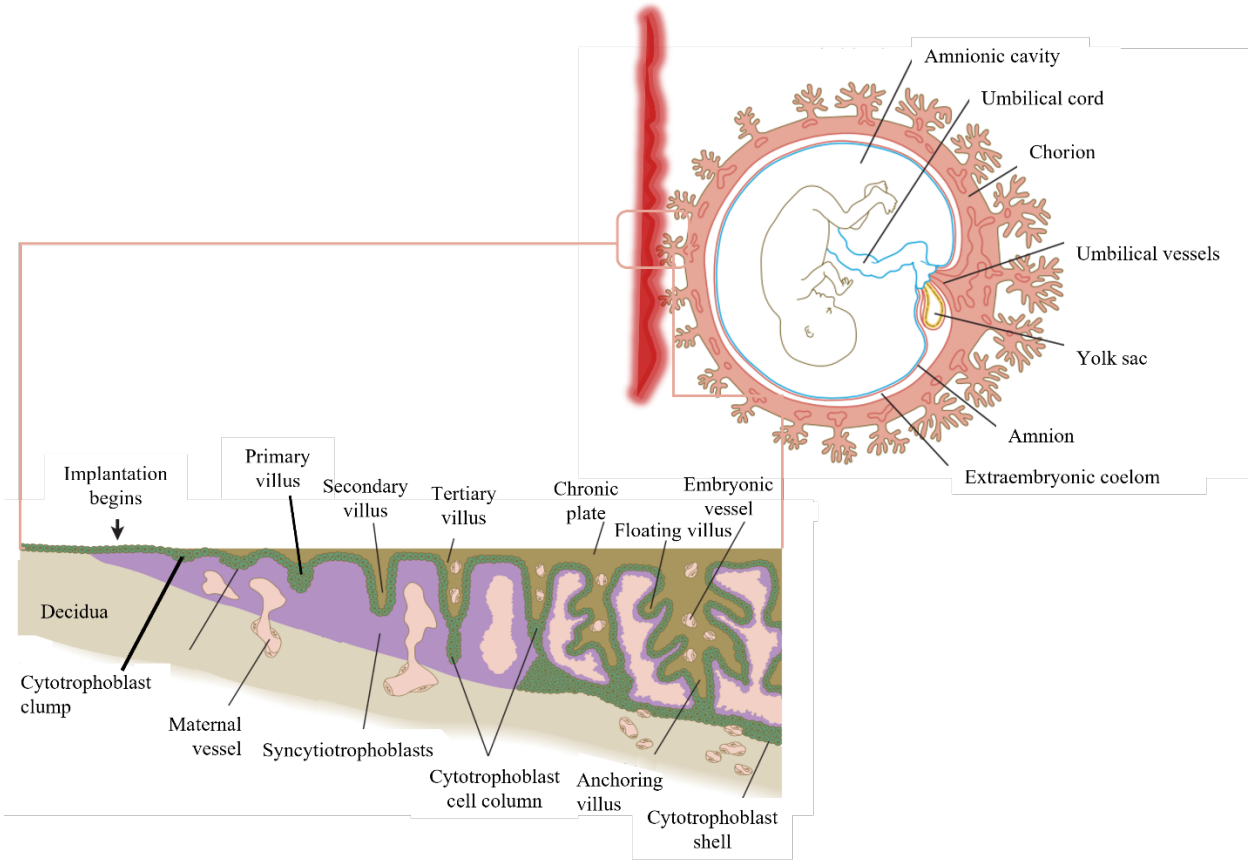


Figure 8: An illustration of the structure and development of the human placenta. The top panel represents the entire structure of the placenta as well as the surrounding extraembryonic features. The placenta is the maternal-fetal interface and the fetal end attached to the fetus by the umbilical cord, and the maternal end is embedded in the decidua. The bottom panel represents a zoomed-in villous tree. It explains the development of its structure according to the steps of primary, secondary, and tertiary villi and its different constituents those stages. This figure was adapted from Human Embryology and Developmental Biology [173].

2.2.2. Preeclampsia (PE)

Preeclampsia (PE) is a complex pregnancy syndrome that affects 2-8% of all pregnancies worldwide, and it is one of the leading causes of maternal, fetal, and neonatal mortality, especially in underdeveloped countries [174, 175]. To this day, PE has only been observed in humans [176]. The etiology of PE remains poorly understood, and therefore pre-diagnosis and prevention are still not feasible [175]. Following the 20th week of gestation, PE is diagnosed by a novel onset of

hypertension (> 140/90 mm Hg) in combination with one or more of the following: proteinuria (>0.3 g/24 h), maternal organ dysfunction (including renal, hepatic, hematological, or neurological complications), or fetal growth restriction [177]. Moreover, severe PE is a major cause of maternal morbidity (i.e. stroke) and adverse perinatal outcomes, such as prematurity and intrauterine growth restriction (IUGR) [175]. PE is subclassified into two types: early-onset PE (EOPE) and late-onset PE (LOPE), depending on whether the onset of clinical symptoms was before or after the 34th week of gestation, respectively[178]. If left undiagnosed and untreated, PE progresses into eclampsia which is the occurrence of seizures superimposed on the syndrome of PE [179]. Most deaths due to PE are avoidable by providing timely and effective care to pregnant women presenting these clinical manifestations. Commonly pregnant women who are at high risk of PE are treated with acetylsalicylic acid (aspirin, 75 mg) before the 20th week of gestation, and patients diagnosed with PE, depending on the severity of their symptoms, are treated with antihypertensive and/or anticonvulsive drugs to alleviate the clinical manifestations in the mother while avoiding unnecessary prematurity and maximizing maternal and infant survival [175, 180].

High-risk PE pregnancies have been described for women with obesity, preexisting hypertension, chronic kidney disease, diabetes, autoimmune disorders, twin pregnancies, pregnancies over the age of 40, a family history of PE, and PE in previous pregnancies [181, 182]. Furthermore, the risk of PE increases with nulliparous pregnancies, *in vitro* fertilization, and change in paternity [183-185]. The only definitive treatment for PE is delivery of the fetus and placenta, after which the mother's health is usually stabilized within 24 hours post-delivery [180]. However, after delivery, the women who have suffered from a PE pregnancy have later in life an increased risk of high blood pressure, heart disease, and stroke [157].

The placenta seems to be a prerequisite for PE. PE was observed in some cases of hydatidiform mole where the uterus contains only disordered placental tissue and no fetus, suggesting that PE is a disease of the placenta [186].

2.2.2.1. The three-stage model of PE

The pathogenesis of PE is described in a three-stage model, summarized in **Figure 9** [187]. Stage 1 involves poor maternal immunoregulation to paternal antigens. In mice, it has been demonstrated that immune priming by coitus induces regulatory T cells (T-regs) because seminal

plasma contains paternal type I and type II MHC antigens and high concentrations of transforming growth factor-beta (TGFβ) [188]. This exposure induces tolerance to paternal alloantigens and an accumulation of uterine T-regs that are important as they facilitate implantation [189]. Changes in coitus partners, infrequent exposure to the paternal seminal fluid, first pregnancies with the specific partner, and artificial insemination by a donor are frequently associated with PE [187, 190]. This suggests that the maternal immune system fails to tolerate paternal alloantigens and may not yet be well adapted to accommodate the semi-allograft fetus. Many researchers have attempted to detect the immunoregulatory activity of the decidua in peripheral blood. They have analyzed the expression of CD4 or CD25 to compare the level of T-reg cells in PE and healthy pregnancies. A healthy pregnancy is supposed to provide an immunotolerant environment for the conceptus, with an immune response that is shifted towards T helper type 2 (Th2), whereas in PE, there is a shift towards a proinflammatory type 1 immunity state. This leads to an increase in pro-inflammatory cytokines, such as IFN-γ, TNF-α, IL-2 that suppress the maternal tolerance system to the fetus [191].

Stage 2 of PE is termed uteroplacental syndrome. In this step, poor placentation occurs due to reduced EVT invasion and impaired maternal spiral artery remodeling as described in **Figure 10**, leading to insufficient placental perfusion causing ischemia and hypoxia. A link has been established between paternal-specificity, maternal adaptation, and successful placentation. The invading EVTs express the well-studied class I molecule HLA-C, which signals paternal specificity that both T-reg cells and decidual natural killer (NK) cells can recognize [187]. NK cells recognize HLA-C via killer immunoglobulin-like receptors (KIRs), resulting in the production of chemokines and angiogenic cytokine, promoting EVT invasion [187]. However, there are two haplotypes of KIRs, and EVT invasion promotion by NK cells is only stimulated by KIR haplotype B while reduced by haplotype A [187, 192]. HLA-C is also highly polymorphic. HLA-C2 interacts with KIRs more strongly than HLA-C1. The combination of HLA-C2 interacting with KIR type B could therefore be the best combination for promoting adequate placentation and protecting against PE, whereas KIR type A mothers presented with HLA-C2 fetuses are the most susceptible to PE [187, 193]. When an inadequate placentation occurs, several pro-inflammatory and anti-angiogenic factors are produced by the placenta. The hypoxic placenta has been shown to stimulate persistently elevated levels of the transcription factor hypoxia-inducible factor 1α and 2 α (HIF1α and HIF2α, respectively) [194]. HIF1α and HIF2α regulate

the expression of hypoxia-induced genes, including erythropoietin, vascular endothelial growth factor (VEGF), nitric oxide (NO) synthase, soluble FLT1 (sFLT1), and soluble endoglin (sEng) [194]. In addition, it has been shown that there is an increase in reactive oxygen species (ROS) in the placenta after repetitive ischemia-reperfusion due to contractile spiral arteries and dysregulation of the haem oxygenase (HO) pathway, a mediator of oxidative stress [194]. The haem oxygenase isoform 1 (HO1) protein, which is substantially lower in the PE placenta, localizes on the perivascular contractile sheath of placental vessels, preventing tumor necrosis factor (TNFα) mediated cell damage [194, 195].

Stage 3 is the clinical stage of PE, and it leads to the maternal syndrome characterized by a new onset of hypertension, proteinuria, and other end-organ damage. Autopsies of deceased PE-patient have revealed pathological endothelial lesions in organs such as the brain and the liver [196]. As PE is a complex and heterogenous disorder involving many factors and pathways, the disease's etiology has not yet been determined. Indeed there may be multiple origins of the diseases, creating multiple subtypes of PE. However, there are some pathways in the disease that have been relatively well characterized so far. In terms of the placenta-produced circulating factors, the most well studied are the pro- and anti-angiogenic factors. sFLT1 binds to the proteins VEGF and placental growth factor (PlGF), antagonizing their proangiogenic signaling pathway [197]. High levels of circulating sFLT1 and low levels of circulating VEGF and PlGF produce an antiangiogenic state that produces systematic endothelial dysfunction and contributes to the clinical manifestations of PE. Another overexpressed antiangiogenic protein in PE women is sEng, which inhibits TGFβ, interferes with endothelial cell stability, and leads to periventricular edema development [198]. Together the very high overexpression of sFLT1 and sEng are suspected of inducing the more severe forms of PE. Interestingly, abnormal levels of these factors are detected about 2-3 months before the onset of PE [198]. The sFLT1-PlGF ratio is currently the standard too for predicting the women which may suffer from PE [199]. PE-induced hypertension is related to the upregulation of angiotensin II type 1 receptor (AT1-AAs) [200]. The AT1-AA agonists are angiotensin II (AngII), antiangiogenic factors, and PE-patient’s production of activating autoantibodies. They induce intracellular ROS production and NF-kB and AP-1 expression, both of which lead to the upregulation of AT1-AA [201, 202]. Zhou et al. demonstrated that immunoglobulin isolated from PE women increased systolic blood pressure four days after injection into pregnant mice, and this hypertensive response was attenuated by the administration

of an AT1 receptor antagonist [203]. Furthermore, it has been shown that the levels of AT1-AA remains elevated even postpartum and may contribute to the increased long-term risk of cardiovascular disease that is observed in women who have suffered from PE [204]. Due to the blend of circulating stress-induced placental factors, the most important resulting clinical manifestations of PE are hypertension (>140/90 mmHg), proteinuria (>300 mg/L protein in 24h), thrombocytopenia (platelet count <100,000), renal insufficiencies, impaired liver function, pulmonary oedema, and visual disturbances.

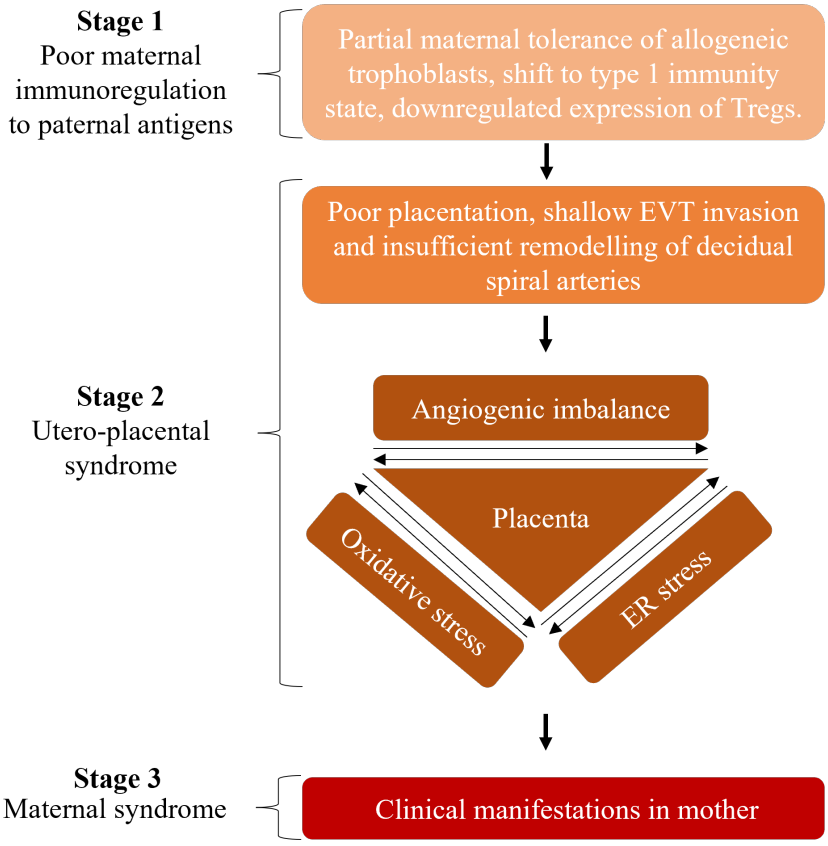


Figure 9: Summary of PE pathogenesis. PE is described in a 3-stage model, each defined by stepwise characteristics that originate from maternal maladaptation to paternal alloantigens, followed by poor placentation that releases harmful factors in the mother's circulatory system, resulting in the clinical manifestations of PE.

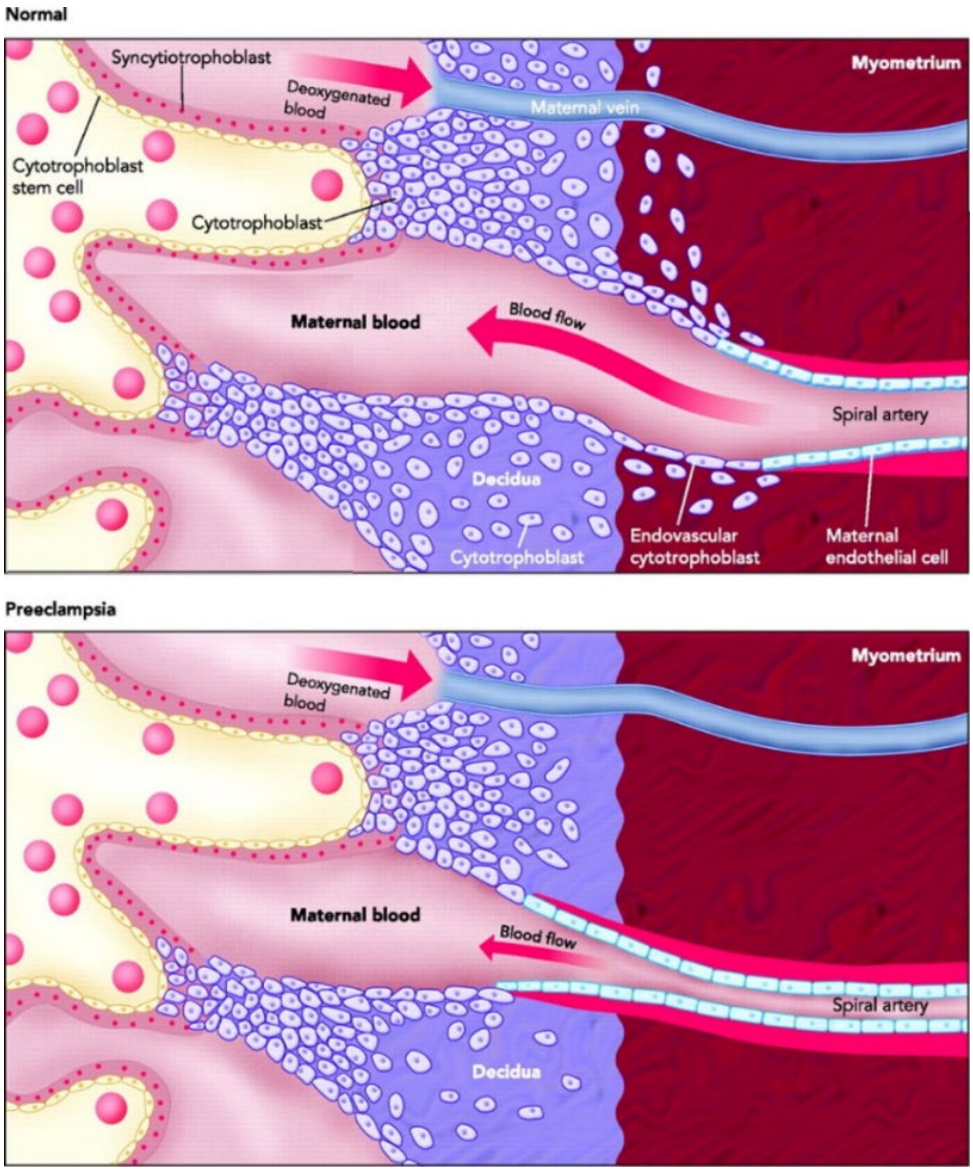


Figure 10: An illustration of placentation difference in a healthy pregnancy versus PE. In normal placental development (top), placental EVTs invade the maternal spiral arteries, transforming them from small to high-caliber capacitance, capable of providing adequate placental perfusion. In a PE placenta (bottom), EVTs invasion of the spiral arteries is shallow, and they remain small caliber, resistant vessels. The image was modified from Wang et al. 2009 [205].

2.2.2.2. The role of DLX5 in placenta development and PE

Distal-less homeobox 5 (DLX5) is a transcription factor belonging to the family of at least six Distal-less homeobox proteins, which are expressed in humans and mice. In humans, during development, DLX5 is expressed in the brain, neural crest cells, ectoderm, skeletal tissue, hematopoietic stem cells, and the trophoblast cells of the placenta [206]. Single-cell sequencing of human embryos revealed that DLX5 is one of the earliest expressed genes in the human trophectoderm of a pre-implantation embryo [127]. The early expression of DLX5 is a human-specific placentation feature demonstrated by comparing DLX5 expression between human, macaque, and mouse embryos, in which macaque displayed a later shifted expression, while the mouse pre-implantation embryo or placenta did not express DLX5 at all [127]. Zadora et al. also showed that DLX5 gene in the human placenta and brain is imprinted, and due to disturbed imprinting of PE placentae leading to a leaky expression, DLX5 is overexpressed in 70% of PE placentae [127]. The upregulation of DLX5 in the placenta affects genes that are associated with cell growth, proliferation, metabolism, survival, and movement [127]. Overexpression of DLX5 in SGHPL4 cells, a human first trimester EVT cell line, generated a transcriptome that clustered with the transcriptome of primary early onset PE patients and is therefore used in research as a PE-model cell line [127].

2.3. Objectives of the work

Our research group has previously deep RNA sequenced 5 control versus 5 EOPE primary human trophoblast samples. The evolutionary youngest, human-specific LINE-1 transcripts, L1HS, are highly expressed in cytotrophoblasts and significantly overexpressed in one PE patient exhibiting exaggerated epigenetic dysregulation [127]. As LINE-1 elements are usually highly suppressed in somatic cells, except in neuro-progenitor stem cells, adult brain cells, and cancer, it is crucial to characterize LINE-1 expression and activity in the trophoblast cells towards understanding if LINE-1 has a role in placenta development. Until now, PE has only been observed in human pregnancies, suggesting that it may be a human-specific disorder; therefore, it is important to understand if the human-specific LINE-1 elements may play a role in PE pathogenesis. In my research, I focused on characterizing the expression and activity of LINE-1 elements in the human trophoblasts and placenta during pregnancy and PE pathogenesis. In addition, as miRNAs are hypothesized to have originated and diversified by TE activity, I also investigated if miRNA dysregulation is observed in PE vs. control placenta samples.

The main objectives of this research were i) to localize in which placenta associated cells are LINE-1s expressed ii) to study the expression level of LINE-1 elements in healthy placentae during the different stages of pregnancy as well as the difference in expression between healthy and PE human placentae iii) to unravel if the human SHG-PL4 EVT cell line can support LINE-1 activity and how is the LINE-1 behavior in the PE model cell line versus wildtype, iv) to study the expression levels of known LINE-1 activators and repressors in healthy and PE placentae, v) to investigate the consequences PE-relevant molecular pathways due to LINE-1 overexpression in the human SGHPL4 cells, vii) to identify if TE-derived and other miRNAs in PE vs. control placenta samples are dysregulated and to identify and validate their target mRNAs in a network of a particular mechanism in order to assess the potential of these miRNAs and genes as potential early biomarkers of PE occurrence.

3. MATERIALS AND METHODS

3.1. Patients Cohorts

3.1.1. Oslo Cohort II

Placenta (Control N=28, EOPE N= 26, LOPE N =24) and decidua (Control N=23, EOPE N=13, LOPE N= 8) samples were obtained from a bio-bank collection at the Oslo University Hospital, Norway, authorized by the Regional Committee of Medical Research Ethics in Eastern Norway, comprising of placentae following C-section. Controls were patients with normal blood pressure and uncomplicated pregnancies undergoing cesarean section due to breech presentation or other reasons.

3.1.2. Berlin Cohort

In addition to placentae samples collected from patients in Zadora et al. study [127], another 3 control and 5 EOPE placentae were obtained from HELIOS Klinikum in Berlin with the approval of the Regional Committee of the Medical Faculty of Charité Berlin, under the same conditions as previously published 5 healthy and 5 EOPE samples[127]. All samples were collected and processed within 2 hours after a Caesarean section delivery. Control patients were defined as having an uncomplicated pregnancy at term. PE patients were defined by hypertension (SBP ≥140 mmHg or DBP ≥90 mmHg) and proteinuria (≥0.3 g in a 24-hour urine specimen) before the 34th week of pregnancy.

3.1.3. Charité Cohort

Samples from 19 placentas were collected at Charité Universitätsmedizin in Berlin. The local ethics committee approved the trial protocol, and consent was obtained from all patients.PE was defined as a new onset of hypertension (140/90 mmHg at two occasions six hours apart), in combination with proteinuria (>300 mg/24 h or >2+ dip stick). IUGR was defined as an estimated fetal weight 95th percentile. Samples from six EOPE placentae, seven placentae from early-onset IUGR complicated pregnancies, and six age-matched control placentae from pregnancies lacking any criteria for PE or IUGR were obtained. Control patients delivered

preterm due to preterm rupture of membranes, preterm labor, and clinical suspicion of chorioamnionitis or poor cardiotocography indicating severe fetal stress.

3.2. Isolation of primary trophoblasts

Table 1 describes all the buffers and chemicals used in this procedure. The placentae were used in primary trophoblast isolation were from the Berlin Cohort. The placentae were collected within two hours of a Cesarian section delivery. The whole placenta was manually processed in ice-cold 0.9% NaCl buffer. The placenta was dissected from decidua parietalis, decidua basalis, and the umbilical cord. The villous tissue was cut into medium chunks and washed with the 0.9% NaCl buffer. Fatty tissue, capillaries, small veins, and fibrotic tissue were removed, and villous tissue was further washed and cut into smaller pieces until pale pink homogenous tissue was obtained. About 100g of tissue was mashed and transferred into an Erlenmeyer flask. The tissue was then digested of enzymatic using trypsin (Sigma Aldrich) and DNase (Roche) at 37°C while shaking. The major cell types were separated using a Percoll (GE Healthcare) gradient (20%-70%), and the cytotrophoblasts were collected between the 3rd and 4th rings from the bottom of the gradient. All further steps were carried out under sterile conditions. Cells were counted using a hemocytometer, and cell viability was assessed by TrypanBlue staining. Cytotrophoblasts were isolated based on the negative selection of the primary monoclonal mouse anti-human HLA-ABC antibody Clone W6/32 (Dako) using Dynabeads Pan Mouse IgG (Thermo Fisher Scientific). The cells/beads solution was placed into a magnet for 3 minutes, and the supernatant was transferred to a new falcon tube, centrifuged at 1000 rpm for 5 minutes, and the pellet was resuspended in 30 ml of SmGM medium (Lonza). Trophoblasts were counted with viability assessment and plated in a cell culture dish for an overnight culture in SmGM medium. The next day, isolated cytotrophoblasts cell pellets were collected and were stored at -80°C for DNA, RNA, and protein isolation.

Table 1: List of buffers and their constituents used in primary trophoblast isolation.

Buffer:	**Recipe:**			
HBSS/HEPES (25 mM HEPES)	- 100 ml 10x HBSS (Hanks Balanced Salt w/o Phenol red, with Ca, Mg) (Gibco) - 5.958 g HEPES (Gibco) - 900 mL ddH_2O - set pH on 7,4 with NaOH			
DMEM/HEPES (25mM HEPES)	- 500ml DMEM (Gibco) - 2.9788g HEPES - Set up pH on 7,4 with NaOH - Sterile filtrate			
Buffer 1A	- 285mg Trypsin (Sigma) - 20mg DNase I (Roche) - 225mL warm HBSS/HEPES			
Buffer 1B	- 10mg DNase I - 25mL warm HBSS/HEPES			
Buffer 2	- 190mg Trypsin - 10mg DNase I - 150mL warm HBSS/HEPES			
Buffer 3	- 150mg Trypsin - 10mg DNase I - 110mL warm HBSS/HEPES			
Buffer 4	- 10mg DNase I - 100mL DMEM/HEPES			
Newborn Calf Serum	- 45mL NBCS (Biochrome) - 7,5mL ddH_2O			
HBSS/HEPES for Percoll	- 10x HBSS w/o Ca, Mg and 250mM HEPES			
90% Percoll	- 5mL HBSS/HEPES - 45mL Percoll (GE Healthcare)			
Percoll gradient	Dilution of Percoll (%)	g/mL	90% Percoll (mL)	HBSS/HEPES (mL)
	70 60 50 40 30 20	1.084 1.071 1.06 1.049 1.038 1.027	3.107 2.667 2.227 1.773 1.333 0.893	0.893 1.333 1.773 2.227 2.667 3.107

3.3. Cell Culture

SGHPL-4 cells are a first trimester EVT cell line. They were a gift from Judith E. Cartwright (St George's University of London, London, United Kingdom). DLX5-over-expressing-SGHPL4-cells (PE-model cell line) was generated as published previously [127]. Both cell lines were cultivated in HAM's F10 (Biochrom) media containing 10% (v/v) FCS, 2mM glutamine and 1% (v/v) penicillin/streptomycin (P/S), at 5% CO2 and 37°C. Table 2 describes the standard materials used in routine cell culture experiments.

Table 2: Standard material used in routine cell culture experiments

Material Name:	**Provider:**
DPBS (without Ca2+ and Mg2+)	PAN Biotech
0.5% Trypsin EDTA	Life Technologies
Fetal Calf Serum (FCS)	Gibco

3.3.1. Retrotransposition Assay

WT SGHPL4 cells and the PE model cell line were used to assay LINE-1 retrotransposition. Each cell line was grown to 70% confluency. In each experiment, cells were seeded in 5 wells in a 6-well plate, $1x10^5$ cells/well in 10% (v/v) FCS and 1% AA Ham's F10 medium (Biochrom) and allowed to attach for 3h in 37°C, 5% CO2. The cells were transfected, using the electroporation Neon Transfection System (Life Technologies) and using the Neon 100µl kit (Invitrogen), with 3 µg of L1-EGFP vector ($pL1_{RP}$EGFP vector), AmaxaGFP plasmid (transfection efficiency and GFP positive control), L1-EGFP mutant vector (pJM111 vector; as a retrotransposition negative control), in individual wells. The pJM111 vector encodes a LINE-1 element with two missense mutations in ORF1 that abolish retrotransposition [51]. All vector maps, except for pJM111, are represented in **Figure 11**. The fourth additional well was used for a mock electroporation control containing no DNA plasmid, and the fifth well contained only WT cells (untreated control). The cells were imaged by fluorescence microscopy using the EVOS FL Cell Imaging System, 7 days post-transfection. Following the imaging, the cells were harvested with trypsin and collected in 500µl PBS. The GFP signal was quantified by the BD LSR Fortessa Flow Cytometer at the MDC Flow Cytometry Technology Platform. The experiment was repeated at least three times. The data was analyzed using FlowJo software.

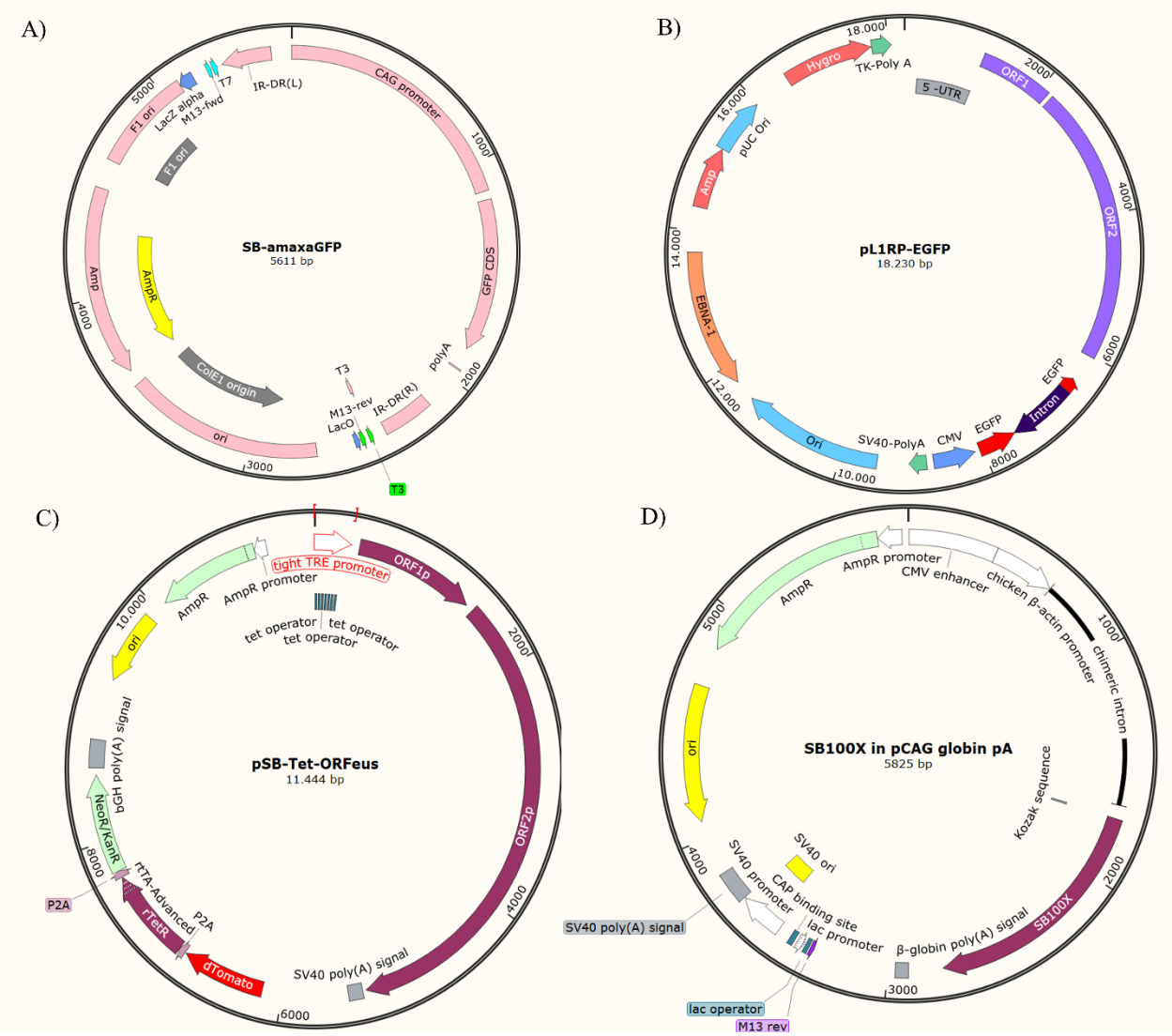


Figure 11: Maps of vectors used for transfection. A) The SB-Amaxa-GFP plasmid is 5611 bp in length, contains an ampicillin resistance gene, a CAG promoter, and a GFP coding sequence. **B)** The pL1RP-EGFP vector is 18230 bp in length, contains an ampicillin resistance gene, the endogenous L1Hs 5'UTR, the LINE-1 ORF 1 and 2 coding sequences, and an EGFP coding sequence including an interfering intron which is spliced out upon a retrotransposition event. **C)** The pSB-Tet-ORFeus vector is 11444 bp in length, contains an ampicillin resistance gene, a Sleeping Beauty Transposase recognition sequence, a codon-optimized human-specific LINE-1 sequence, with a doxycycline-inducible promoter. **D)** The SB100X vector is 5825 bp long and contains an ampicillin resistance gene and a coding sequence for the Sleeping Beauty transposase. All vector maps are displayed in the software SnapGene Viewer.

3.3.2. LINE-1 Overexpression in SGHPL4 Cells

SGHPL4 cells at confluence ~70% were washed twice with PBS without Ca2+ and Mg2+ and harvested with trypsin. Cells were collected in Ham´s F10 medium, 10% FBS, counted, and centrifuged at 500 x g for 5 min. After two subsequent PBS washing steps, cells were resuspended in 60µl Resuspension Buffer R (Neon 100 µL kit, Invitrogen) to a total of $1.4x10^5$ SGHPL4 cells. Cells were electroporated with 10:1 ratio of the vector carrying *L1ORFeus* overexpression cassette (5 µg DNA) and plasmid pcGlobin2-*SB100X* containing hyperactive *Sleeping Beauty* transposase (500 ng DNA) with Neon Transfection System (Life Technologies) using parameters as follow: 1260V/20ms/2 as pulse voltage/pulse width/pulse number. After electroporation, cells were cultured in a 6-well plate in 2 ml Ham´s F10 medium containing 10% FBS. The experiment was repeated at least three times. All vector maps, except for the mutate L1 vector (negative control) are represented in **Figure 11**. L1 overexpression was confirmed by RT-PCR.

3.3.3. Transwell invasion assay

Transwell invasion assay was performed on WT SGHPL4 cells subjected to human recombinant chemoattractants EGF, POSTN, and IGF1. The transwell inserts were pre-chilled and then coated with 50 µl of growth factor reduced matrigel (0. 25 mg/ml) diluted in Hams F10 medium w/o FBS. The plate containing the inserts was then incubated for 24 hours in a cell culture incubator at 37°C. The SGHPL4 were serum-starved by incubating them for 24 hours in Ham's F10 media containing only 0.5% FBS. The transwell inserts were prepared in two steps. In the first step, an equal number of serum-starved SGHPL4 were plated onto the Matrigel-coated wells. Serum starved cells were rinsed with 5 ml DPBS (without Ca2+ and Mg2+), harvested in 1 ml trypsin, and collected in 4 ml of Ham's F10 media with 0% FBS. Cells were counted by hemocytometer and centrifuged at 300 x g for 5 minutes. After removing the supernatant, cells were resuspended in 0% FBS Ham's F10 to obtain 5x105 cells /ml cell suspension. Then, 200 µl of cells ($5 X 10^4$) per well were plated onto the matrigel coated inserts. In the second step, chemoattractants, listed in **Table 3**, were prepared. EGF (10ng/ml) was used as a positive control and was prepared in 5% FBS in Ham's F10, and as a negative control, 5% FBS in Ham's F10 was used. The rest of the wells were filled with 750 µl of 5% FBS + Ham's F10 per well, containing

either 100ng of POSTN or IGF1, 100ng of POSTN and IGF1, 500ng of POSTN or IGF1, and 500ng of POSTN and IGF1. Cell invasion lasted on average 16 hours at 37°C. The invaded cells were fixed with 3.7% paraformaldehyde (PFA) while incubated for 2 minutes at room temperature. Cells were rewashed 2X with PBS to remove the PFA completely. 100% methanol was added for permeabilization, and the cells were incubated for 20 minutes at room temperature. The transwell inserts were washed twice and the cells both invaded and non-invaded were stained with 0.2% crystal violet. The plate containing the cells was covered with aluminum foil and then incubated for 15-20 minutes at room temperature. Cells were washed again after incubation. Non-invaded cells were scraped off by using cotton swabs. Images were taken under the EVOS FL Cell Imaging System (bright field). The colorimetric quantification was done by dissolving the crystal violet from the invaded end of the transwell inserts in 0.1% SDS, while shaking for 1 hour at 37°C. The colorimetric values were measured at the absorbance of 570nm on a Tecan Microplate reader - Spark.

Table 3: List of chemoattractants used in the transwell invasion assay.

Material name:	**Provider:**
Recombinant Human Epidermal Growth Factor (hEGF)	PeproTech
Recombinant Human (hPOSTN)	PeproTech
Recombinant Human (hIGF1)	PeproTech

3.4. RNA isolation and analysis

3.4.1. Total RNA Isolation:

Total RNA was isolated using Trizol lysis reagent and Direct-zol™ RNA MiniPrep kit, including DNase I on-column digestion (Zymo Research), according to the manufacturer's protocol. The RNA concentration was quantified on NanoDrop Spectrophotometer ND-1000, and the quality of RNA was analyzed using RNA 6000 Nano Kit (Agilent) on the Aligent 2100 Bioanalyzer.

3.4.2. miRNA isolation

The miRNAs were isolated from the Oslo cohort II patient placentae, using the mirVana™ miRNA isolation kit (Thermo Fisher Scientific) according to the manufacturer's protocol. The isolated miRNA samples were cleaned up from DNA with a DNase digestion (Zymo Research). The concentration of RNA was quantified on NanoDrop Spectrophotometer ND-1000, and the quality of RNA was analyzed using RNA 6000 Nano Kit (Agilent) on the Aligent 2100 Bioanalyzer.

3.4.3. Reverse Transcription (cDNA generation)

RNA was reverse transcribed into cDNA using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) according to the manufacturer's protocol. For miRNA expression detection, the RNA was reverse transcribed into cDNA the Exiqon miRCURY LNA™ First-strand cDNA kit, according to the manufacturer's protocol.

3.4.4. RNA Sequencing Library Preparation and Data Analysis

The library for RNA sequencing was prepared from isolated RNA using Illumina TruSeq Stranded mRNA LT Set A kit, according to the TruSeq Stranded mRNA Sample Prep LS Protocol. All samples were indexed with sample-specific indices, which allowed for the pooling and

sequencing of all libraries. 150 bp first strand-specific paired-end reads were used, and the BGI Group performed the high-throughput sequencing as a service.

The RNA-sequencing data analysis was done by Dr. Amit Pande from Dr. Zsuzsanna Izsvak's research group. In short, the following is the description of the bioinformatic RNA sequencing analysis. Raw reads filtering software tools such as the FASTX-Toolkit and Trimmomatic were used to discard low-quality reads, trim adaptor sequences, and eliminate poor-quality bases. Outliers with over 30 % disagreement were discarded. For read alignment, Salmon was used to build index and align the reads using the following commands: salmon index -t transcripts.fa -i transcripts_index --decoys decoys.txt -k 31, ./bin/salmon quant -i transcripts_index/ -l IU -1 fastq -2 fastq --validateMappings –o output. Quantified data was checked for GC content and gene length biases using R package NOISeq to provide useful plots for quality control of count data. Mean variance and PCA were computed between biological replicates using the tximport package in R using lengthscaledTPM (CPM cutoff >2 and sample cutoff 2 between the replicates) for the analyzed groups (High LINE-1 PE and High LINE-1 control). Batch effects were removed using the RUv package of Bioconductor Subsequently, the samples were normalized using TMM (weighted trimmed mean of M-values) method for differential expression analysis. The differential expression analysis was conducted using the DESeq2 package. TE identification and quantization across samples was employed using genomic coordinates obtained from Repeatmasker package. The TE sequences were concatenated in the RNA seq data using the cat command. Qauntification (RPKM) of all samples with respect to TE expression was performed using SEAL package (Sequence Expression AnaLyzer) comprising of BBDuk http://jgi.doe.gov/data-and-tools/bb-tools/ employing the following command: seal.sh in=rep1.fastq in2=rep2.fastq ref=TE.transcriptome.fa stats=TE.sealstats.txt rpkm=TE.rpkm.txt ambig=random.

3.4.5. RT-PCR

The real-time polymerase chain reaction (RT-PCR) was conducted on ABI 7900HT Fast Real Time PCR System (Applied Biosystems), using the program outlined in **Table 4**. The data was analyzed by the Sequence Detection Systems 2.4.1 software (Applied Biosystems). Primers used for genes or genetic elements are listed in **Table 5**. All primers used for genes or genetic

elements were checked according to the general guidelines for primer design using Oligo Calc: Oligonucleotide Properties Calculator http://biotools.nubic.northwestern.edu/OligoCalc.html. The primers were synthesized by BioTez Berlin-Buch GmbH. The primer pairs used for miRNA detection were purchased from Qiagen, formerly Exiqon, are hsa-miR-190b-3p, hsa-miR-190b-5p, hsa-miR-103a-3p, hsa-miR-210-3p, hsa-miR-576-5p, hsa-miR-135-5p, hsa-miR-31-5p. All miRNA primer pairs contained the locked nucleic acid (LNA) technology. The expression of all genes or genetic elements was normalized to 18S expression. The expression of all miRNAs was normalized to hsa-miR-103a-3p expression. All miRNA cDNA was diluted 1/40 prior to RT-PCR. For the RT-PCR reactions, 2x Power SYBR Green PCR Master Mix (Applied Biosystems) was used according to the manufacturer's protocol.

Table 4: Program steps of the real-time polymerase chain reaction.

Step	**Stage 1 – Pre-denaturation**	**Stage 2 - Denaturation**	**Stage 3 - Annealing**		**Stage 4 - Extension**		
Temperature	50°C	95°C	95°C	60°C	95°C	60°C	95°C
Time	2 min	10 min	15 s	1 min	15 s	15 s	15 s

Table 5: Primers used for genes and genomic element expression in RT-PCR experiments.

Gene/Fragment Name	**Forward Primer (5' to 3')**	**Reverse Primer (5' to 3')**
L1HS	GAATGATTTTGACGAGCTGAGAGAA	GTCCTCCCGTAGCTCAGAGTAATT
LINE-1-ORFeus	GCTGGATGGAGAACGACTTC	TTCAGCTCCATCAGCTCCTT
h18S	ACATCCAAGGAAGGCAGCAG	TTTTCGTCACTACCTCCCCG
hIRF7	TGCAAGGTGTACTGGGAG	TCAAGCTTCTGCTCCAGCTCCATAAG
hIFITM1	ACTCCGTGAAGTCTAGGGACA	TGTCACAGAGCCGAATACCAG
hPOSTN	TGCCCAGCAGTTTTGCCCAT	CGTTGCTCTCCAAACCTCTA
hIGF1	CACACCATGTCCTCCTCGCA	CAGAGCGTCTCCGGTCCAG

3.5. Protein Isolation, Detection, and Analysis

3.5.1. Protein Isolation

Protein was isolated from placenta tissue or cells. Placenta tissue was pulverized in liquid nitrogen. For protein extraction, about 20mg tissue was measured and subjected to homogenization with glass beads. The homogenate or cells were lysed in RIPA Lysis Buffer (the recipe is described in **Table 6**) supplemented with the cOmplete™ Protease Inhibitor Cocktail (Sigma Aldrich) and Benzonase (1:2000, Merck Millipore). The homogenate or cells were resuspended in the supplemented RIPA buffer, cooled on ice for 30 minutes, and then the lysate was centrifuged at 4 °C for 15 minutes at 12000 x g. An aliquot of the lysate was separated for Bradford Assay for total protein quantification, and the rest of the protein lysate was mixed with 6X SDS Loading Buffer and boiled at 95°C, and then the samples were either frozen or used directly for Western Blot.

3.5.2. Western Blot

The Western Blot buffer recipes are described in **Table 7**. Denatured proteins were separated on a 10% SDS-PAGE gel prepared by the TGX stain-free FAST cast acrylamide gel kit (Bio-Rad), according to the manufacturer's protocol. Proteins were then transferred to a Trans-Blot Turbo midi-size PVDF membrane (Bio-Rad) using a semi-dry blotting system (BioRad). Unspecific binding was blocked with 5% w/v skim milk (Fluka) in TBS-T. Primary antibodies: mouse monoclonal anti-LINE-1 ORF1p clone 4H1 (1:500, Merck), chicken polyclonal anti-LINE-1 ORF2p (1:20,000) were incubated overnight at 4°C. The secondary antibodies, goat anti-mouse (1:5000, Thermo Scientific) and rabbit anti-chicken (1:5000, Thermo Scientific), conjugated with horseradish peroxidase, were used. Blotted membranes were detected with Super Signal™ West Femto Maximum Sensitivity Substrate (Thermo Scientific). Membranes were then stripped using mild stripping buffer (200mM glycine, 0.1% SDS v/w, 1% Tween20 v/v, pH 2.2), blocked again and incubated with a primary antibody or milk for storage.

Table 7: List of buffers and their constituents used in protein lysis and Western Blot.

Buffer:	Recipe:
RIPA Buffer	- 0.05 M Tris-HCl pH 7.4 - 0.150 M NaCl - 0.001 M EDTA - 1% Triton-X100 (Sigma) - 1% Na-Deoxycholate - 0.1% SDS (ROTH)
6X SDS Loading Buffer	- 3 mL Glycerol - 1.5 mL β-Mercaptoethanol - 9 mL 10% SDS (ROTH) - 1 M Tris-HCl (pH 6.8) - Pinch of Bromophenol Blue - Adjust to 10 mL ddH_2O
1X Transfer Buffer	- 200mL 5X TransBlot Turbo transfer buffer - 600 mL ddH_2O - 200 mL Ethanol
10X TBS (pH 7.5)	- 0.5 M Tris - 1.5 M NaCl - Adjust to 1L with ddH_2O
1 X TBS-T	- 100mL 10X TBS - 0.1% Tween 20 (Calbiochem)
Mild Stripping Buffer	- 0.2M Glycine - 0.1% SDS - 1% Tween-20
Ponceau S Staining Solution	- 0.1% Ponceau S (Sigma) - 5% Acetic Acid

3.5.3. Immunohistochemistry Staining

Human control and PE placenta sections were deparaffinized and dehydrated. Antigen was retrieved by heat-induced epitope retrieval method in 0.01 M Citrate buffer, pH 6. Endogenous peroxidase activity was blocked with hydrogen peroxide (Lab Vision/Thermo scientific) for 10 minutes. Tris-buffered saline with 0.05 % Tween-20 (TBS-T) was used in all 14 washing steps. Non-specific immunoglobulin binding was blocked with Ultra V Block (Lab Vision). Mouse anti-LINE1-ORF1p clone 4H1 antibody (Merck) was diluted (1:500) in PBS and was incubated overnight at 4°C. To identify non-specific tissue staining, antibody diluent without the primary antibody was used as a negative control. The secondary antibody Peroxidase-conjugated goat anti-mouse IgG (Jackson ImmunoResearch) was diluted (1:100) in PBS and was incubated for 30

minutes at room temperature. Peroxidase activity was detected with aminoethylcarbazole (AEC) chromogen (Thermo Scientific) for 10 minutes. For nuclei staining, tissue sections were counterstained with Haematoxylin and mounted with Kaiser's glycerine gelatine (Merck). Stained placenta tissue slides were analyzed by Dr. Martin Gauster from the Institute for Cell biology, Histology and Embryology of Medical University of Graz, Austria.

3.5.4. Immunofluorescence Staining

WT-SGHPL4 cells were seeded in a 8-well chamber slide (ibdi) and incubated in Ham´s F10 + 10% + 1% AA overnight at 37°C, 5% CO_2. The cells were washed twice with PBS, then were fixed using 4% paraformaldehyde (Sigma) for 15 minutes and permeabilized with 0.1 X-100 in PBS for 2 minutes. The cells were blocked with 5% milk for 1 hour. The chambers were then incubated with the primary antibody mouse monoclonal anti-LINE-1 ORF1p clone 4H1 (1:500, Merck) and incubated overnight at 4°C with gentle shaking. After washing, the secondary antibody donkey Alexa Fluor 647 anti-mouse (1:200, Invitrogen) was incubated for an additional 1 hour. After extensive washes with PBS, the nuclei were stained with DAPI solution (Vectashield with DAPI) for 10 minutes at room temperature in the dark. After additional washing, the samples were mounted using ProLong® Gold Antifade Reagent (Invitrogen).

3.6. miRNA target prediction

For miRNA-targets prediction, TargetScan 7.2, the online available search tool for predicted miRNA taergets in mammals, was used. The human species was human, and the search query was the miRNA miR-576-5p. In short, TargetScan predicts biological targets of miRNAs by searching for the presence of conserved 8mer, 7mer, and 6mer sites that match the seed region of each miRNA [207]. In mammals, predictions are ranked based on the predicted efficacy of targeting as calculated using cumulative weighted context++ scores of the sites [208]. The TargetScan target gene output for miR-576-5p for *POSTN* and *IGF1* genes is shown in **Table 8.**

Table 8: TargetScan output for query has-miR-576-5p and the targets *POSTN and IGF1*

Target gene	Site counts				6mer sites	Cumulative weighted context++ score
	total	8-mer	7-mer-m8	7-mer-A1		
POSTN	1	0	0	1	0	-0.01
IGF1	2	1	1	0	4	-0.01

3.7. Previously Analyzed or Online-Available Data

3.7.1. Oslo Cohort II Microarray Data Analysis

The bioinformatics analyses of the Oslo Cohort II microarray data (consisting of 50 PE patients including EOPE (N = 26) and LOPE (N= 24) and 28 controls) was previously analyzed by Dr. Manvendra Singh, a former bioinformatician in Dr. Zsuzsanna Izsvak's, and the analysis is briefly described here. The R platform (https://cran.r-project.org/) and various Bioconductor packages (https://bioconductor.org/biocLite.R) were employed to analyze the human placenta and decidua microarray data. The significance level was extracted from normalized expression values against the background corresponding to each probe IDs using "*lumi*", a R Bioconductor package where the variance-stabilizing transformation (VST) was applied to deal with sample replicates and robust spline normalization (RSN) for normalization.155 Probes with p-value < 0.05 were further transformed to log2 scale, and IDs were annotated as gene names from "*illuminaHumanV3.db*" from Bioconductor annotation data package containing 47,324 probes. An expression matrix of 38,382 significant probes was generated. The expression values of a gene with multiple probes were assigned by their mean, resulting in 26,886 unique genes. Variation from upper and lower quantiles of resulting expression sets was corrected by quantile normalization; validation was confirmed by Principal Component Analysis (PCA). Each gene value was further assigned as their relative abundance value, which is the ratio of gene expression value in each sample and the mean value of expression for the corresponding gene across the samples. The resulting relative expression matrix was subjected to unsupervised hierarchical clustering, which blindly classifies all samples based on their transcriptome pattern without the prior knowledge of their disease status (Spearman correlation and distances between observations were calculated using Euclidian distances and average linkage). P-value threshold for the correlation test of matrix was kept up to 0.01. To examine the statistical reliability of clustering, bootstrapping (1000 replicates) on unbiased hierarchically clustered dendrogram of distance matrix by Ward method using pvclust was performed. In the present study, differential gene expression fold change between samples on the $\log_2$ scale was analyzed using linear and e-Bayesian model algorithms from "*limma*" R Bioconductor package.

3.7.2. Embryo Development Single-cell RNA Sequencing Data

Raw single-cell RNA-seq datasets from early human embryos and embryonic stem cells (GSE36552), and the EPI, PE, TE cells (GSE66507) were downloaded in the sra format. Following the conversion of raw files into fastq format, the quality was determined by using the FastQC. Two nucleotides were removed from the ends as their quality scores were highly variable compared with the sequences in RNA-seq reads. Before aligning the resulting reads, the reference genome annotations using the "Hot L1" classification first curated, that are ~90 full-length LINE-1 sequences capable of coding both ORFs. The genes (genecode V19) and LINE-1 subgroups genomic sequences were extracted, and they were combined to generate a reference transcriptome. These sequences were then appended, comprising the coding-sequences plus UTRs of genes and locus-level LINE-1 subgroups sequences in fasta format. Each fasta sequences was annotated with their respective genes or LINE-1 IDs. To guide the transcriptome assembly, each of the resulting contigs was appended and modelled was in gtf format that was utilized for the expression quantification. Next, the concatenated genes and LINE-1 subgroups transcriptome and genome reference sequences were indexed using 'salmon'. Finally, the trimmed sequencing reads were aligned against the curated reference genome. The 'salmon' tool quantified the counts and normalized expression (Transcripts per million (TPM)) for each single-cell RNAseq sample. Overall, this approach enabled to simultaneously calculate LINE-1 subgroups and protein-coding gene expression using expected maximization algorithms. Data integration of obtained count matrix, normalization at logarithmic scale, and scaling was performed as per the "Seurat V.2.6" (http://satijalab.org/seurat/) guidelines. The annotations of cell-types were taken as it was classified in original studies. Differential expression was calculated, and their level of significance was tested using Kruskal–Wallis test by comparing cell types of interest with the rest of the cells. The obtained p-values were further adjusted by Benjamini-Hochberg method to calculate the False Discovery Rate (FDR). All the statistics and visualization of RNA-seq was performed on R (https://www.r-project.org/). This data was analyzed by Dr. Manvendra Singh, a bioinformatician in Dr. Zsuzsanna Izsvak's research group.

3.7.3. Genome-wide screen for LINE-1 regulators

Liu et al. [96] provided the first genome-wide list of genes involved in LINE-1 retrotransposition control, using CRISPR/Cas9 screening in two distinct human cell lines, K562 and HeLa cells. This list of LINE-1 regulators is used in this study and displayed in **Table 9**. In short, the cells already contained a BFP-Cas9 lentiviral transgene and were transfected with the pB-tetO-L1-G418R/Blast construct and the piggyBac transposase. The Cas9/L1-G418R cells were lentivirally infected with a genome-wide sgRNA library, containing ~200,000 sgRNAs targeting 20,549 protein-coding genes and 13,500 negative control sgRNAs. Following doxycycline induction, genomic DNA was extracted. The sgRNA-encoding constructs were PCR-amplified using Agilent Herculase II Fusion DNA Polymerase and the libraries were then sequenced across by Illumina NextSeq. Computational analysis of genome-wide screen was performed using Cas9 high-throughput maximum Likelihood Estimator (casTLE). For each gene, casTLE combines measurements from multiple targeting reagents to estimate a maximum effect size as well as a p-value associated with that effect.

Table 9: List of LINE-1 regulators provided by Liu et al [96] expressed in the human placenta.

Gene (Activator)	Gene (Repressor)
AK4	CXCL11
ALDH4A1	DEPDC5
ALDOA	ERCC1
ALG3	EZH2
AP1S2	FAN1
APEX1	HTR1F
APOBEC3A	INO80
APOBEC3F	MECP2
APOBEC3G	MLLT10
BAZ1B	MORC3
CCDC130	MPHOSPH8
CGA	OPRK1
CPSF7	PLA2G12A
DCLRE1C	RBM10
DHX9	RDH13
DLG4	RNASEH2B
FANCE	RNF166
FCGR1B	RPRD2
MEN1	SAMHD1
MRPL28	SCAMP3
NSMAF	SETDB1
RAB36	SETX
RCL1	SLX4
SLC11A2	ZMYND11
SNRNP25	
SRSF11	
SUV39H1	
XPNPEP1	
ZFHX2	

3.7.4. Small RNA Sequencing Data

Along with microarray, small RNA sequencing was also performed on the same set of PE and controls of the Oslo Cohort II patient samples. The bioinformatics analysis of this data was analyzed by Dr. Manvendra Singh from Dr. Zsuzsanna Izsvak's, and the analysis is briefly described here. PE samples Sample-specific barcoded sequencing reads were de-multiplexed from multiplexed flow cells. The resulting BCL files were converted to FASTQ format files using CASAVA 1.8.2. The quality of the raw sequence reads was determined with the FastQC.[56] Reads with a quality score < 30 were removed. Two nucleotides were truncated from the end of sequencing reads since their average quality score was not the same as the rest of the nucleotides. This resulted in at least 0.10 million reads per sample. Next, the reads were mapped over the reference genome (Human hg19/GRCh37) and transcriptome model (hg19.refseq.gtf), downloaded from USCS tables (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/) using Bowite 2.0.5.0 applying parameters as: *"—very-sensitive-local"*. On average 75% of total reads were mapped on the reference genome. For calculation of relative abundance of miRNAs, Counts Per Million (CPM) were calculated using feature counts and algorithms from "*DESeq2*" which performed quantization and statistical inference of systematic changes between conditions, as compared to within-condition variability. For the above, the reference miRNA sequences were provided from miRbase database. Once the relative and normalized counts were obtained, the data processing was performed similarly with the above microarray data. Overall, the expression matrix of genes from microarray and expression matrix of miRNAs from small RNA-seq for the same set of samples was obtained. Their relative to the mean expression values were separately computed across all the samples, and matrix multiplications followed by Spearman's correlations were performed. This provided another matrix where the level of correlation was defined for each gene with each miRNA. Finally, the predicted and validated set of miRNA targets from miRbase were obtained, and those in our correlation matrix were inspected. Heatmaps and the rest of the plots were made using R as written above.

3.8. Chemicals

Chemicals were obtained from the following companies: Karl Roth GmbG & Co. KG, Jena Bioscience, Merck KGaA, Sigma-Aldrich GmbH, Amersham-Pharmacia, Qiagen, GE Healthcare, BD Bioscience and Miltenyi Biotec GmbH.

3.9. Statistics

Statistical analysis was performed with GraphPad Prism 7.0 Software and R Statistical Programming Language. Data are presented as either mean ± SEM (for normally distributed data) or median with inter-quartile range (for non-normally distributed data). Normal distribution was assessed by D'Agostino & Pearson omnibus normality test, ShapiroWilk normality test and Kolmogorov-Smirnov tests. Data sets were compared using the Unpaired t-test, Mann–Whitney test, Wilcoxon-rank sum test, ANOVA, and Kruskal-Wallis test as appropriate. Post hoc testing included Dunnet's test and Dunn's test for multigroup comparisons. Statistical correlation analysis between nonparametric variables was performed using Spearman´s rank correlation. The techniques for each analysis are specified in the figure legends. Two-tailed testing with a normal-based 95% confidence interval was performed for each analysis and $p < 0.05$ was considered statistically significant.

4. RESULTS

4.1. Evolutionary youngest and human-specific LINE-1, L1Hs, are highly expressed in the trophectoderm lineage in embryo development and remain specifically and highly expressed in cytotrophoblasts of the placenta.

To understand the expression dynamics of the human-specific LINE-1 elements during human embryo development, online available single-cell RNA sequencing data of early human embryos were analyzed for the time point L1HS expression and cell type, shown in **Figure 12A**. L1HS is already expressed at the 2-cells, 4-cell stages, and the expression level peaks at the 8-cell stage. Following the formation of the blastocyst, the TrE is the only cell lineage that expresses L1HS elements compared to the cells of the ICM (inner cell mass), EPI (epiblast), and PEndo (primitive endoderm), and this approximately 2-fold difference in expression is statistically significant (q value = 0.03647). In **Figure 12B**, the transcript expression of L1HS in the human placenta CTBs was confirmed by a RT-PCR in a pregnancy-related tissue panel containing samples of the CTBs, decidua, macrophages, muscle, fat, and HUVEC cells (umbilical cord cells). The data shows a tendency towards a higher expression in the CTB versus macrophages, significantly higher expression (p-value = 0.0251) in CTBS versus the human umbilical vein endothelial cells (HUVEC). Macrophages and HUVECs can be more easily separated from CTBs, therefore is no possibility of L1HS contamination, unlike in the placenta, decidua, and fat, where different trophoblast cells are embedded.

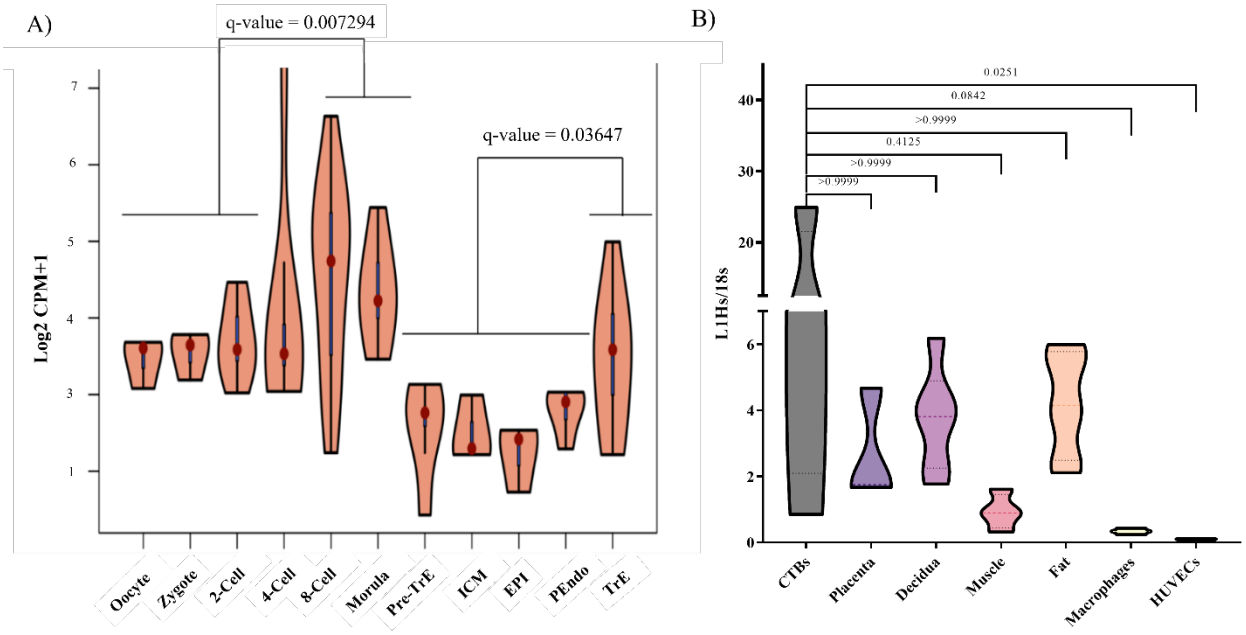


Figure 12: Expression of human-specific LINE-1 elements, L1HS, in human embryo development and in pregnancy associated tissue. A) A violin plot of single-cell RNA sequencing data showing the distribution of normalized expression profiles of L1Hs in early human embryonic cells, and human embryonic stem cells. Abbreviations: Pre-TrE (pre trophectoderm), ICM (inner cell mass), EPI (epiblast), and PEndo (primitive endoderm), TrE (trophectoderm). **B)** A violin plot of RT-PCR data expressing human L1HS normalized to 18S. Abbreviations: CTBs (cytotrophoblasts), HUVEC (human umbilical vein endothelial cells). The statistics were calculated using the Kruskal–Wallis test: Dunn's multiple comparisons test.

4.2. L1HS are highly expressed in the human cytotrophoblast cells without significant difference in expression in placentae samples between EOPE-patients and controls from three different cohorts.

To define which LINE-1 families are transcriptionally expressed in human trophoblast cells, deep RNA sequencing was conducted on isolated primary human trophoblast cells from healthy and PE placentae. **Figure 13A** shows the transcription levels of the seven youngest evolutionary LINE-1 families that have invaded the ancestral genomes. Results of this transcriptome analysis have revealed that only the evolutionary youngest LINE-1 elements, L1HS, are expressed in human trophoblasts versus the non-human specific and older LINE-1 families. The data has also

revealed that in three out of ten PE patients' CTBs (PE1, PE16, and PE25), L1HS is significantly higher expressed (p-value = 0.01) in the patients versus all others, as shown in **Figure 13B**.

To determine if L1HS expression differs in PE patient trophoblasts versus healthy placentae in a larger sample cohort, RT-PCR for L1HS was conducted in two different cohorts containing healthy and PE placentae samples as shown in **Figure 13C** and **D**. L1HS was not significantly dysregulated in EOPE or LOPE versus Control in the Oslo Cohort II placenta samples, and was found to be downregulated in LOPE versus the gestationally matched controls (p value = 0.0256). Since PE is a disorder that originates in the first trimester of pregnancy, first trimester placentae were investigated for L1HS expression in PE low risk (LR) (N=9) versus high risk (HR) (N=10) patient placenta samples. The PE risk was evaluated based on the patient's medical history. **Figure 13E** shows no significant difference in expression in LR versus HR PE samples, except for one patient in the HR group expressing very high L1HS versus all other patients.

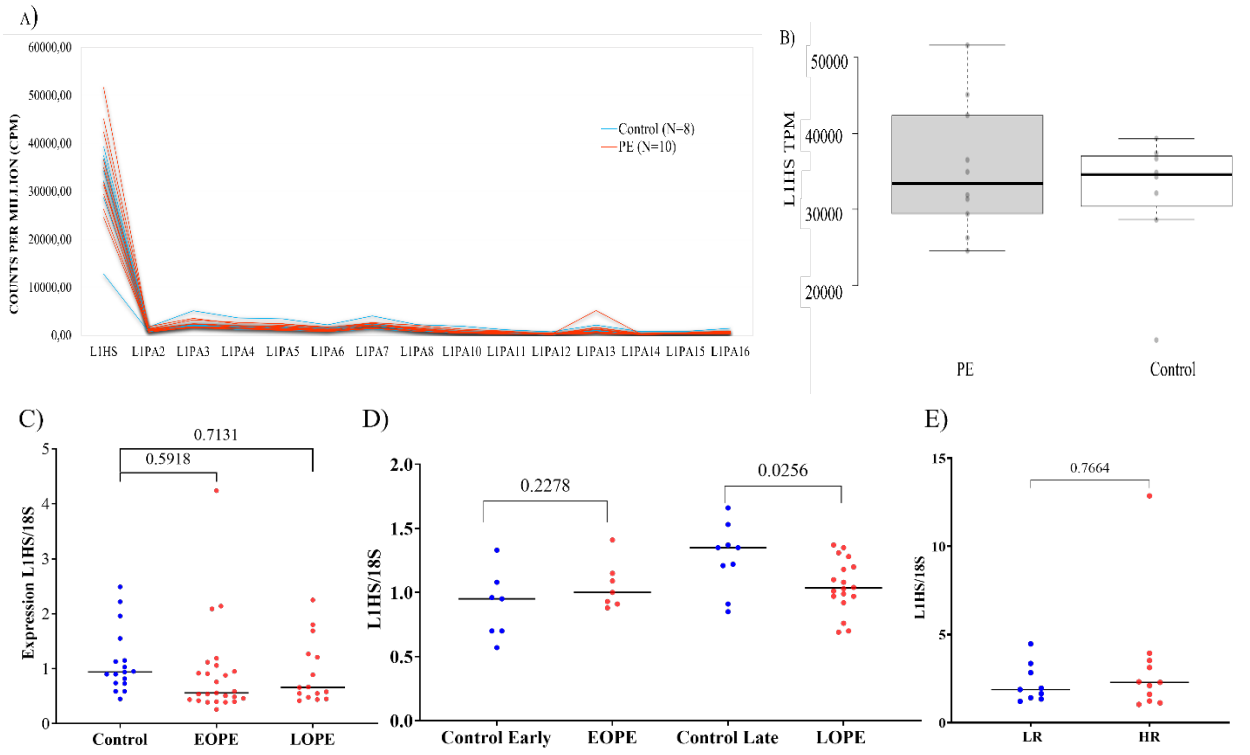


Figure 13: LINE1 RNA expression levels in Control versus PE patients across different cohorts. A) A line graph of the RNA expression levels of all LINE-1 subfamilies from RNA sequencing data of control (N=8) versus EOPE (N=10) isolated trophoblasts from the Berlin Cohort placentae. **B)** Boxplot of L1HS RNA expression in PE (N=10) vs control (N=8) isolated

trophoblasts of the Berlin Cohort RNA sequencing. P-value is 0.01, and the interquartile range (ICR) confidence interval equals to 95%. **C)** L1HS RT-PCR data on RNA expression levels in control (N=17) versus EOPE (N=19) and LOPE (N=19) placentae samples from Oslo Cohort II. The statistics were calculated with Ordinary One-Way ANOVA. **D)** L1HS RT-PCR data on RNA expression levels were also probed in early or late pregnancy control versus EOPE and LOPE, respectively, placentae samples of the Charité Cohort. For statistical calculation, the unpaired t-test was used. **E)** RT-PRC data of L1HS RNA expression levels in PE low risk (LR) (N=9) versus high risk (HR) (N=11) first-trimester placenta samples. The statistics were calculated using the unpaired t-test.

4.2.1. L1HS correlation with PE-relevant parameters

To better understand the relationship of LINE-1 expression with pregnancy and PE-relevant parameters, the mRNA expression of L1HS was correlatated to clinical data, such as available for the samples from the Oslo II Cohort (third trimester), as shown in **Figure 14**. **Figure 14A** displays that there is no correlation between L1HS expression and placenta efficiency. The placenta efficiency is the ratio of the baby's weight and the placenta weight. In **Figure 14B** it is shown that the sFlt1/ PIGF ratio, the PE prognostic marker that is determined by measuring the levels of PIGF and sFLT1 in patient sera, also does not correlate with the expression of L1HS. Using the published data for DLX5 mRNA expression that was previously RT-PCT quantified and normalized to 18S [127], it was found that DLX5 expression significantly correlates with the expression of L1HS, as shown in **Figure 14C**.

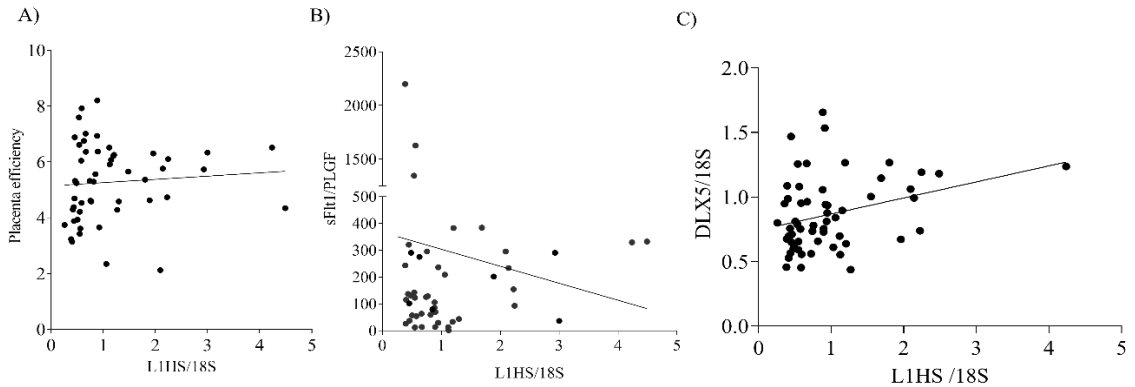


Figure 14: L1HS correlation with Oslo Cohort II clinical and published data determined by Spearman rank correlation. A) Placenta efficiency (baby's weight/placenta weight) versus

L1HS/18S correlation: r = 0.04317; P-value = 0.7056. **B)** sFlt1/ PIGF ratio versus L1HS/18S correlation: r = - 0.1377; P-value = 0.2233. **C)** DLX5/18S versus L1HS/18S correlation: r = 0.2772; P-value = 0.0351.

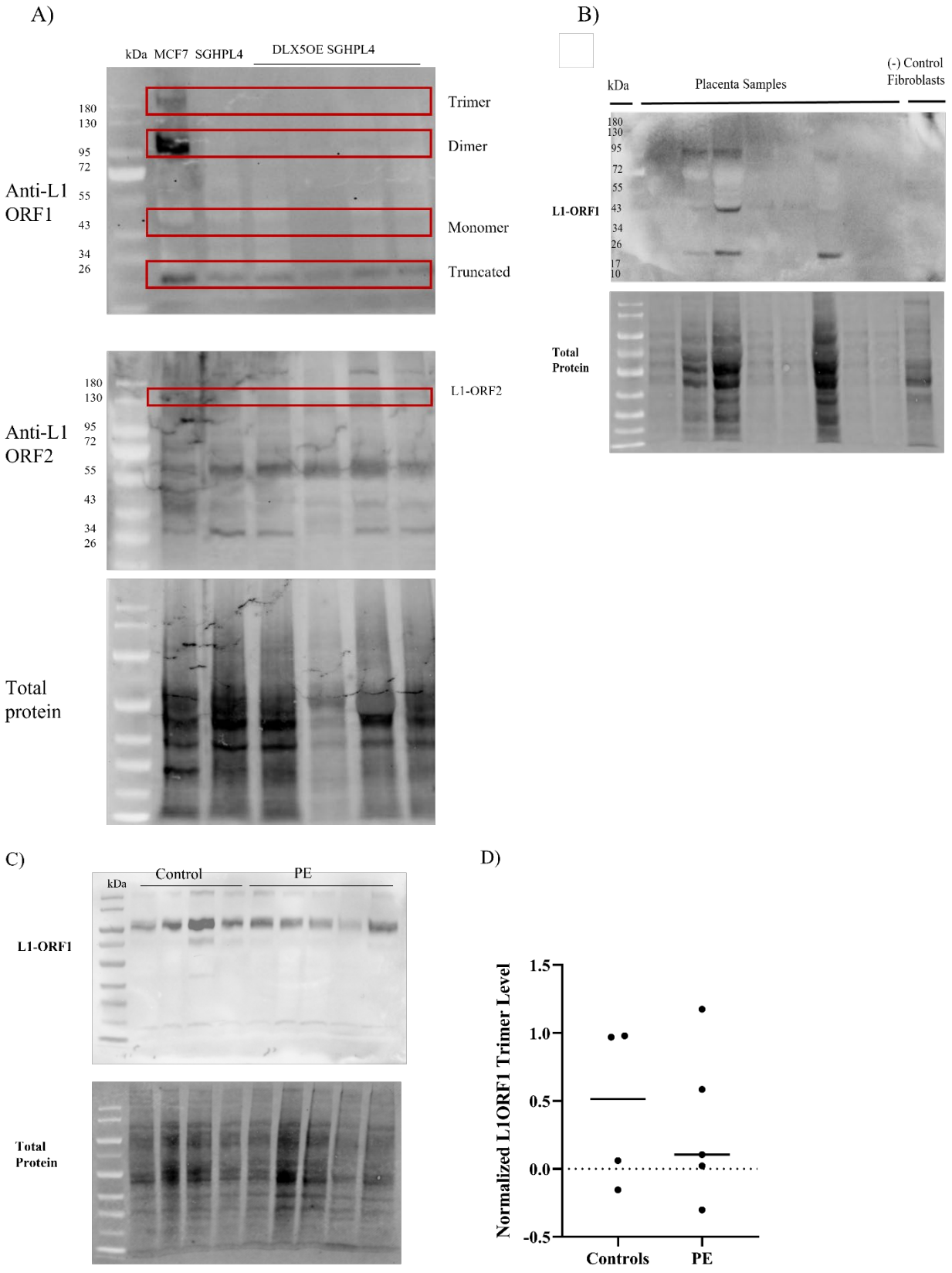
4.2.2. LINE-1 ORF1 and possibly ORF2 are expressed at the protein level in human placentae, in all three trophoblast subtypes, and LINE-1 ORF1 is localized in the cytoplasm and nuclei of the human SGHPL4 cell line.

Next, for LINE-1 elements to propagate, both ORF1 and ORF2 proteins are necessary to be expressed. Western Blot analysis was conducted to verify if both L1-ORF1 and ORF2 proteins are expressed in the human placenta or trophoblast cells. **Figure 15A** shows the detection L1-ORF1(upper panel) and ORF2 (lower panel) in cells lines MCF7 (positive control), SGHPL4, and DLX5OE-SGHPL4. L1-ORF1 protein was detected at 43kDa, the truncated protein, as well as the dimer, and trimer (active form) of the protein. In each well, 50 µg of total protein was loaded. The L1-ORF2 band appears at the expected molecular weight at 150kDA; however, this detection cannot be concluded due to many unspecific bands. **Figure 15B** shows L1-ORF1 protein detection in human placenta samples and a human fibroblast isolated protein as a negative control, showing that no forms of L1ORF1 are present in the fibroblast negative control sample. In contrast, the truncated, monomer, and dimer forms of L1-ORF1 are present across the placenta samples. **Figure 15C** is a Western blot of L1ORF1 protein expression between control and PE patient placenta samples. All three forms of L1ORF1 are present except for the monomer form that is present as a faint band in one control sample. In each well, 50 µg of total protein was loaded. To quantify the level of expression of the L1ORF1 trimer form in control placenta sample versus PE placenta samples, the L1ORF trimer band was compared to the total loaded protein. As shown in **Figure 15D**, there was no significant difference in expression between control versus PE patient placenta L1ORF1 protein levels in the trimer form.

To better understand which subtypes of trophoblast cells of the placenta express LINE-1s, immunohistochemistry (IHC) was conducted. **Figure 15E** displays the IHC controls. Adjacent tissue sections were used to detect L1ORF1 and the EVT marker, HLA-G, to confirm that the expressing cells are indeed EVTs. The negative control for the L1ORF1 antibody was conducted by incubating the secondary antibody without the primary antibody on the tissue to verify the

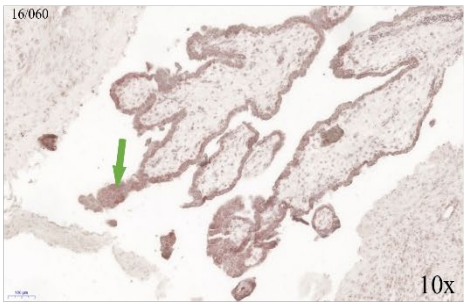
secondary antibody specificity and noise signal. In **Figure 15F**, IHC of L1-ORF1 was conducted on first trimester and third-trimester healthy placentae samples. In this result, L1-ORF1 expression is detected in all three subtypes of trophoblast cells, the STBs and the unlaying CTBs, as well as the invading EVTs, each indicated by a purple, black, and green arrow, respectively. According to the staining intensity of the cells between the first and third trimester, it seems that L1-ORF1 is more highly expressed in the first trimester placentae than in the third.

Lastly, to visualize subcellular localization of L1-ORF1 in the EVTs, immunofluorescence (IF) was performed on SGHPL4 cells. As shown in **Figure 15G**, L1-ORF1 is localized mostly in the cytoplasm of the cells in a punctate manner, as previously published by others [209]. Some L1-ORF1 signals are also displayed in the nucleus of the cells as confirmed by colocalization with DAPI staining.

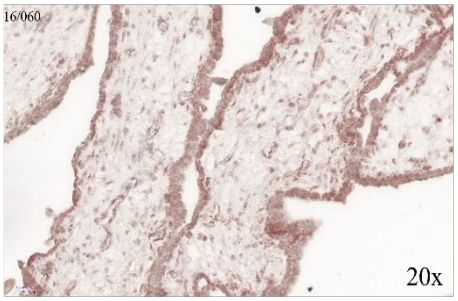

A)
kDa MCF7 SGHPL4
DLX5OE SGHPL4
180
130
95
72
55
43
34
26
Anti-L1 ORF1
Trimer
Dimer
Monomer
Truncated
Anti-L1 ORF2
L1-ORF2
Total protein
B)
kDa
Placenta Samples
(-) Control Fibroblasts
L1-ORF1
Total Protein
C)
kDa
Control
PE
L1-ORF1
Total Protein
D)
Normalized L1ORF1 Trimer Level
1.5
1.0
0.5
0.0
-0.5
Controls
PE

E)

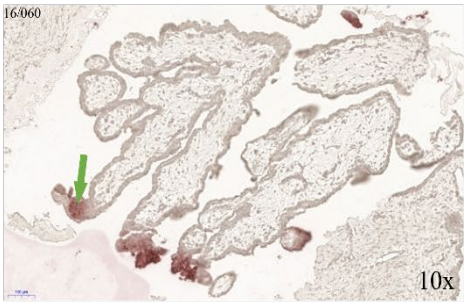
Anti-Line-1 ORF1



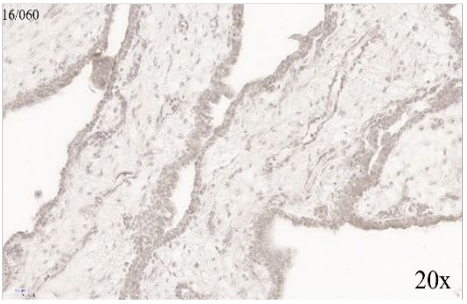
Anti-Line-1 ORF1



Anti-HLA-G (EVT + Control)



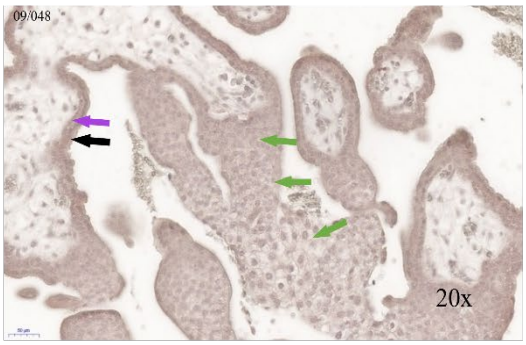
Secondary Antibody Only (- Control)

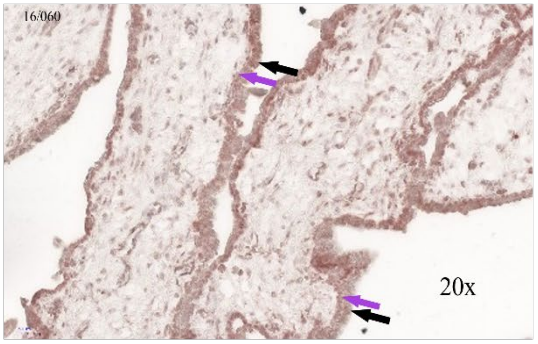


F)

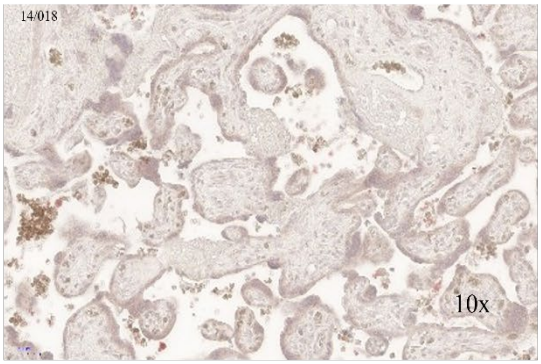
Anti-Line-1 ORF1

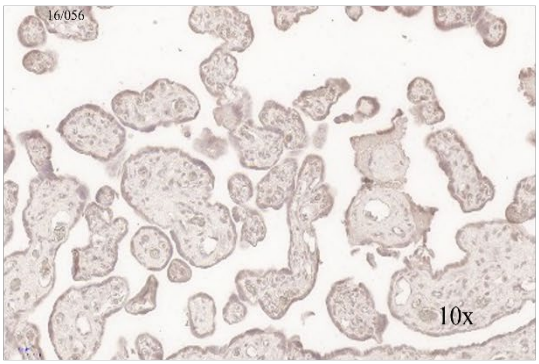
1st trimester





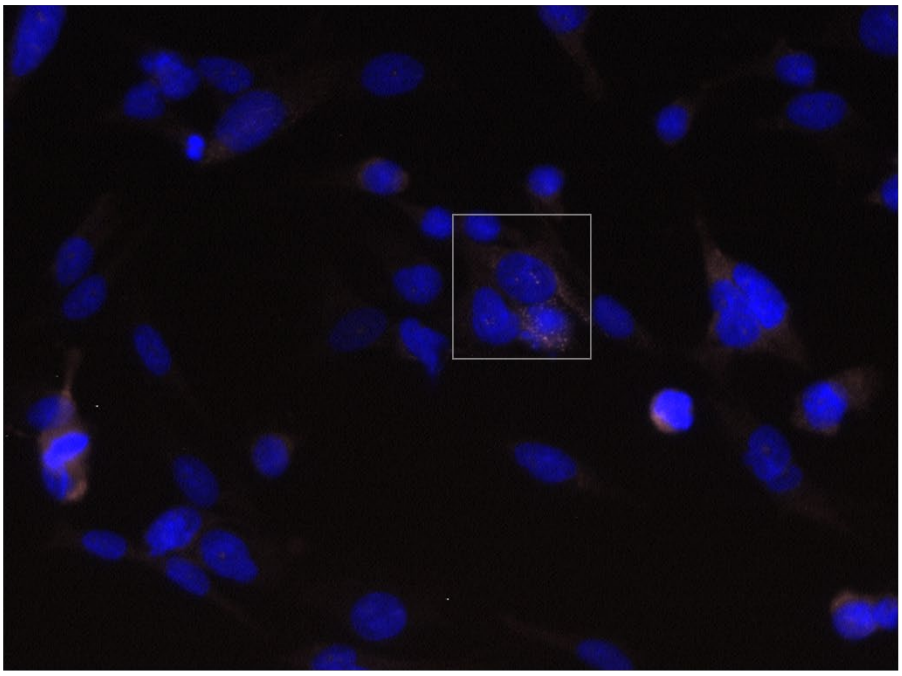
3rd trimester





G)

L1-ORF1, 10X



Secondary antibody only (- Control), 10X

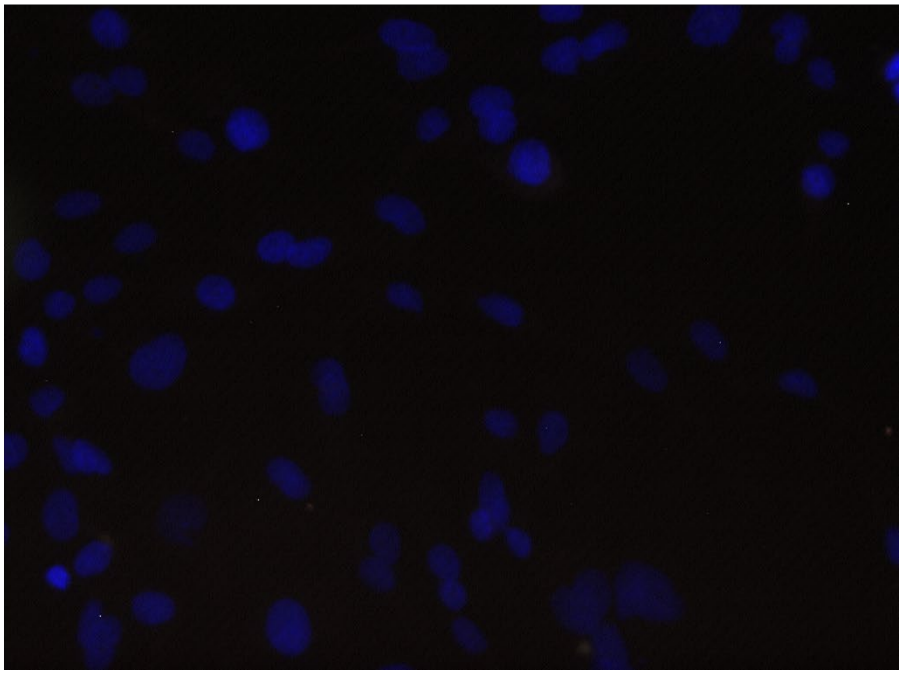


Figure 15: LINE-1 proteins detection by three different methods in cell line and placentae isolated protein samples. A) Western blot of L1ORF1 (upper) and L1ORF2 (lower) expression in MCF7 (positive control), SGHPL4, and DLX5OE-SGHPL4. **B)** Western blot of L1ORF1

expression in human placenta samples and a fibroblast protein sample (negative control) **C)** Western blot of L1ORF1 expression in PE vs control placentae protein samples**. D)** Quantification of the L1ORF1 trimer in PE vs control placentae protein samples. Statistics calculated by unpaired t-test. **E**) IHC staining of L1ORF1 and HLA-G in a first-trimester placenta (right) and secondary antibody only (left) as a negative control. Arrow legend: green = EVTs. **F)** IHC staining of L1ORF1 in a first-trimester placenta (upper) and third-trimester placenta (lower). Arrow legend: green = EVTs, purple = CTBs, black = STBs. **G)** IF staining of WT-SGHPL4 cells with the L1-ORF1 antibody and DAPI (left) and secondary antibody only (right) as a negative control, 10x magnification.

4.3. LINE-1 is retrotranspositionally active in human SGHPL4 cells and overactive in the PE model cell line (DLX5OE-SGHPL4 cells)

A retrotransposition assay was conducted in the trophoblast cell line, SGHPL4 cells (WT) to investigate if trophoblasts expressed all the necessary proteins for LINE-1 retrotransposition and if the cells provide a favorable environment for the activity to take place. In addition, this assay also tested if the PE model cell line would also support retrotransposition, and if the activity would differ from the WT cells. The retrotrasposition assay involved cell transfection with the L1-EGFP reporter or the L1-EGFP mutated reporter (data not shown) and detection of the GFP signal via fluorescence microscopy and Flow Cytometry analysis. The method is illustrated in **Figure 16A**. **Figures 16B-D** show the results of the retrotrasposition assay show the expression of the GFP signal, which indicates that trophoblast cells are expressing all required proteins for retrotransposition and provide a favorable environment for the activity to occur. In comparing WT SGHPL4 cells retrotransposition activity versus the DLX5OE-SGHPL4 cells, DLX5OE-SGHPL4 cells display a significantly higher activity of retrotrasposition based on the GFP signal detected.

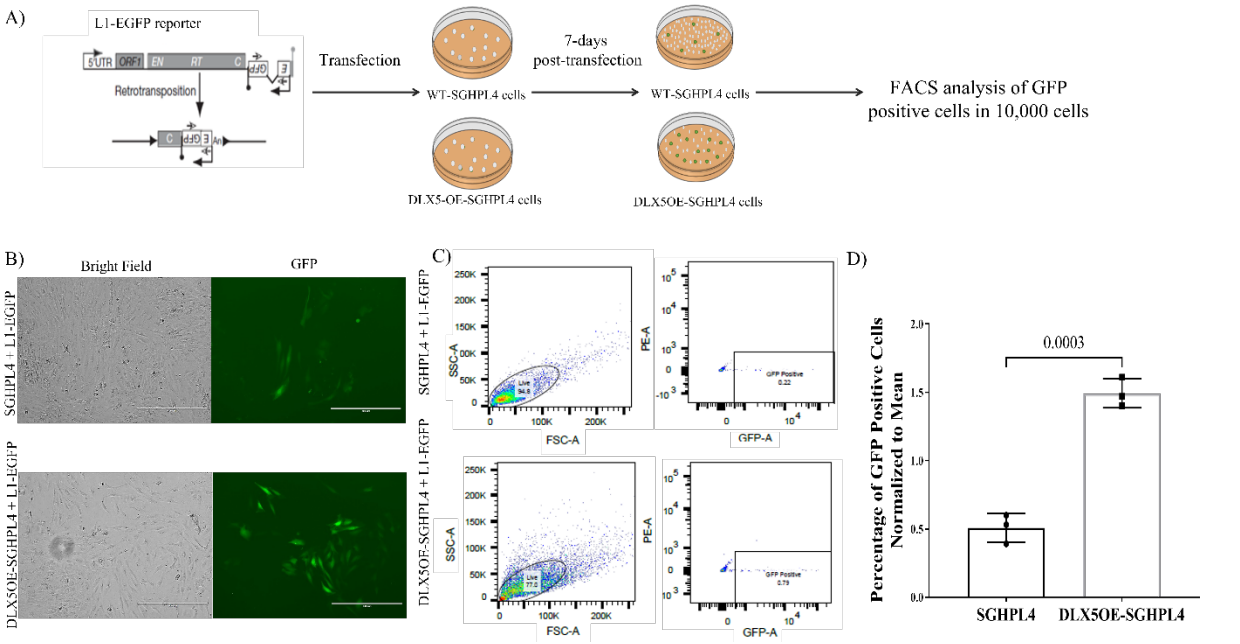
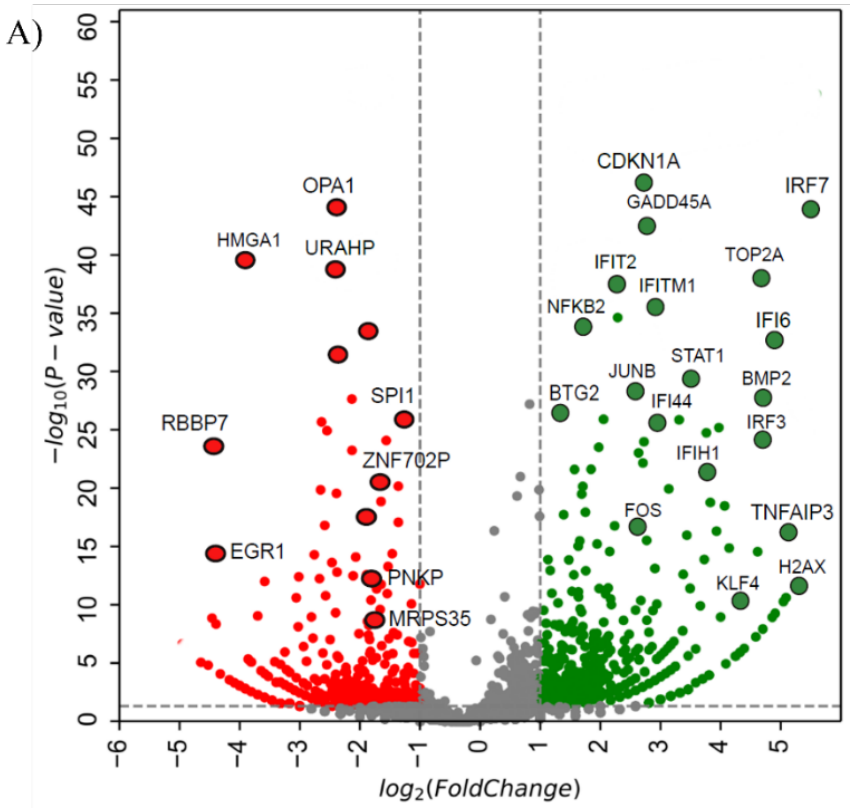


Figure 16: LINE-1 retrotransposition assay in SGHPL4 versus DLX5OE-PL4 cells. A) The rationale behind the retrotransposition assay: A reporter gene cassette, interrupted by a backward intron and inserted in opposite transcriptional orientation into the 3' UTR of the LINE-1, and is expressed only when the LINE-1 transcript is spliced, reverse-transcribed, its cDNA inserted in the genome, and the reporter gene is expressed from its promoter. **B)** Images of SGHPL4 and DLX5OE-SGHPL4 cells 7 days post-transfection under bright field and GFP light acquired by the EVOS FL Cell Imaging System at 10X magnification. **C)** Flow cytometry analysis detecting GFP positive cells in SGHPL4 and DLX5OE-SGHPL4 cells 7 days post-transfection. **D).** Quantification of the GFP signal (indicating a retrotransposition event) in SGHPL4 vs. DLX5OE-SGHPL4 cells. The retrotransposition experiment was repeated three times. In each experiment, the final values were normalized to the mean of both values, and the statistics were calculated using the unpaired t-test.

4.4. LINE-1 overexpression impacts the levels of DNA damage, type I interferon, and cell senescence gene expression in PE versus controls

Based on the available literature, LINE-1 overexpression has been shown to induce the expression of genes involved in the following pathways: DNA damage (such as GADD45A,

ERG1, CDKN1, BTG2, NFKB2), type I interferon (such as IFIH1, IFIT2, IFI6, IFI44, IFITM1, IRF3, IRF7, STAT1), and cell senescence (TNFAIP3, RBBP7, JUNB, FOS, BTG2, KLF4, BMP2, HMGA1, H2AX, OPA1). The expression of 92 genes that are the most important for those pathways was analyzed in three PE patients and three control samples in which L1HS expression was significantly higher versus other patient samples in the Berlin Cohort RNA sequencing data. The results of this analysis, shown in **Figure 17A**, indicate that most of the genes involved in the pathway mentioned above are significantly upregulated in the L1Hs high-expressing patient samples. To verify that L1HS overexpression induces type I interferon pathway genes upregulation, RT-PCR was performed on IRF7 and IFITM1 in L1-ORFeus-OE-SGHPL4 cells. As shown in **Figure 17B**, upon doxycycline induction of L1-ORFeus overexpression, both IRF7 and IFITM1 are upregulated significantly.



B)

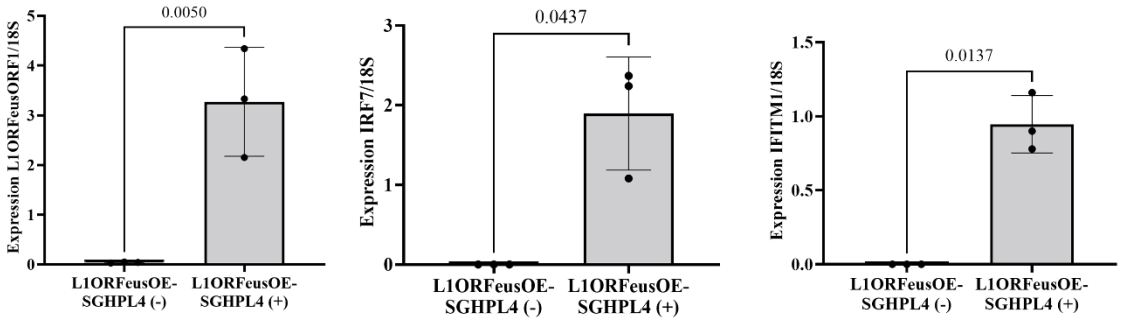


Figure 17: A) A volcano plot of DNA damage, type I interferon, and cell senescence gene expression in L1HS high expressing PE versus L1HS high expressing control primary trophoblast samples from the Berlin Cohort RNA sequencing data. **B)** RT-PCR quantification of L1ORFeus-ORF1, IRF7, and IFITM1 in L1ORFeusOE-SGHPL4 cells. Legend: (-) = non-induced expression of the transfected vector, (+) = doxycycline induced expression of the transfected vector. The statistics were calculated using the paired t-test.

4.5. Dysregulated LINE-1 regulators in PE versus control placentae

To understand how LINE-1 is regulated in the human placenta and potentially in PE, available online data of LINE-1 regulators from Liu et al. [96] was mined for LINE-1 regulators that are expressed in the placenta, based on the Oslo Cohort II expression data. **Figure 18** is a chart of all 54 placetna expressed LINE-1 regulators expression values versus the published Combo casTLE effect of those genes acting either as activators or repressors of LINE-1s. The numeric value of the casTLE effect describes the impact of the given regulator on LINE-1 expression. The positive or negative sign Combo casTLE effect indicates if the regulator is an activator (negative sign) or repressors (positive sign). The 11 genes labeled and highlighted in orange are significantly dysregulated (p-value < 0.05) in PE versus control placentae in Oslo Cohort II microarray data, they are also displayed in **Table 10**. In **Figure 18**, LINE-1 activators with a strong Combo Castle Effect Score (<1) that are not dysregulated in PE versus control placentae of Oslo Cohort II, are SNRNP25, SETDB1, RCL1, DLG4, SRSF11, RBA36, APOBEC3G, and MRPL28. The repressors with a strong Combo casTLE Effect score but not dyregulated in Oslo Cohort II data are MECP2. MEN1, SLX4, DH9, SAMHD1, RNASEH2B, FANCE, ALDOA, MPHOSPH8, ERCC1, RPRD2.

To check if LINE-1 regulators repetoir is differentially expressed in the subgroup fo PE patients that significantly higher levels of L1HS, the LINE-1 regulators expression was also checked in the trophoblast RNA sequencing data in the LINE-1 high-expressing PE patients versus LINE-1 high-expressing controls, and the significantly dysregulated regulators are presented in **Table 11**. The set of regulators that is significantly differentially expressed in the total patients Oslo II cohort versus the select LINE-1 high expressing patients differs. There are two LINE-1 regulators, SAMHD1 (repressor) and RAB36 (activator), that have a strong Combo Castle Effect score and are significantly upregulated in the L1HS high-expressing group.

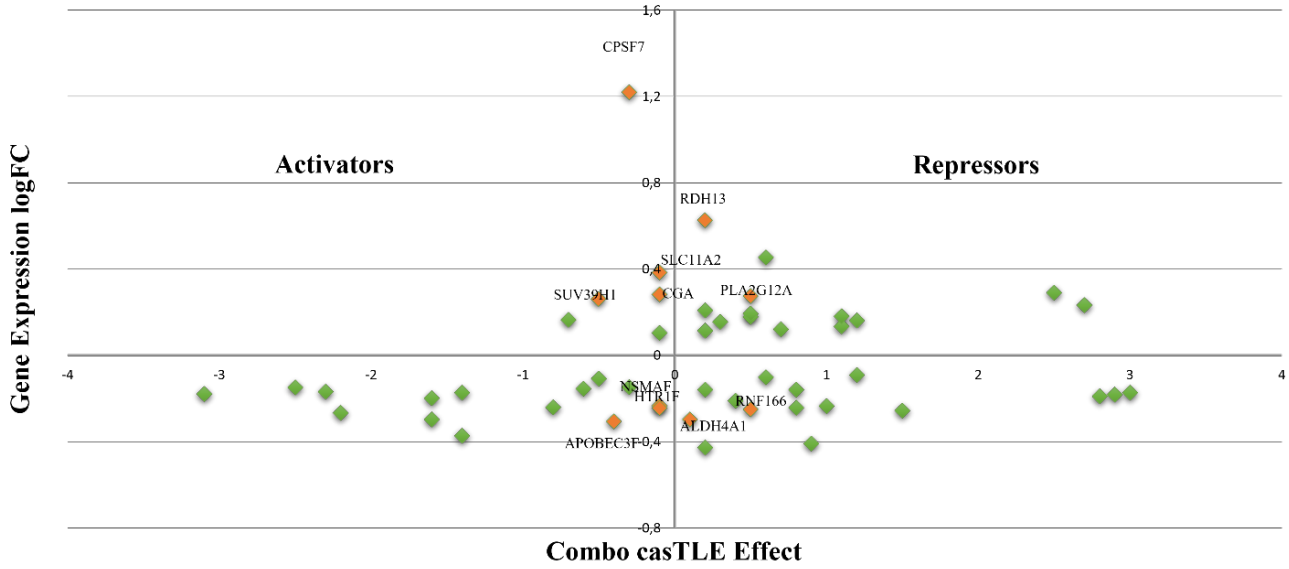


Figure 18: A scatter plot of Oslo Cohort II microarray data logFC versus the regulator combo castle Effect from Liu et al. [96]. Highlighted in orange are genes which are significantly dysregulated in PE vs C patient placentae based on the microarray data analysis.

Table 10: LINE-1 regulators significantly dysregulated in PE (N=26) versus controls (N=23) placentae from the Oslo Cohort II Microarray data.

	LINE-1 Regulator Data [96]		Oslo Cohort II – Placenta Microarray Data	
Gene	**Combo casTLE Effect**	**Regulation**	**Fold Change**	**P-value**
SLC11A2	-0,1	Activator	0,38	0,00024
RDH13	0,2	Repressor	0,63	0,003556
RNF166	0,5	Repressor	-0,25	0,007214
NSMAF	-0,1	Activator	-0,23	0,008746
HTR1F	-0,1	Activator	1,06	0,010624
CPSF7	-0,3	Activator	1,22	0,011676
APOBEC3F	-0,4	Activator	-0,31	0,014465
SUV39H1	-0,5	Activator	0,26	0,023554
PLA2G12A	0,5	Repressor	0,27	0,027871
CGA	-0,1	Activator	0,28	0,029042
ALDH4A1	0,1	Repressor	-0,30	0,034439

Table 11: LINE-1 regulators significantly dysregulated in LINE-1 high-expressing PE (N=3) versus LINE-1 high-expressing controls (N=3) in the trophoblasts RNA sequencing data.

	LINE-1 Regulator Data [96]		Berlin Cohort – Trophoblast RNA Sequencing Data	
Gene	**Combo casTLE Effect**	**Regulation**	**Fold Change**	**P-value**
SAMHD1	1,2	Repressor	3,38	0,00246
OPRK1	0,2	Repressor	2,22	0,00402
RAB36	-1,6	Activator	1,71	0,01326
APOBEC3A	0,4	Repressor	-2,54	0,02943
EZH2	0,5	Repressor	1,22	0,04673

4.6. TE-derived miRNAs and their target genes expression in human placentae and dysregulation in PE.

To decipher which miRNAs are dysregulated in PE versus control placentae previously conducted and analyzed small RNA sequencing data from the Oslo Cohort II was probed. In total, 48 miRNAs are significantly dysregulated (p-value < 0.05) in PE versus control placentae. The top 20 dysregulated miRNAs are shown in **Table 12**. Particularly of interest were the TE-derived miRNAs, according to the data from Qin et al. [138]. There are 3/48 TE-derived miRNAs which are hsa-miR-548b (originating from a DNA transposon) with a log fold change (logFC) value of 0.90 (p-value = 0.000525), hsa-miR-576 (originating from a LINE-1 retrotransposon) with a logFC of 0.45 (p-value = 0.014), and hsa-miR-224 (originating from the DNA transposon MER 135) with a logFC of 0.49 (p-value = 0.0378). As LINE-1 elements are of interest in this research, hsa-miR-576-5p was focused on. Firstly, the small RNA sequencing data analysis was validated by RT-PCR using selected top dysregulated miRNA candidates (has-miR-210, has-miR-31-5p, has-miR-576-5p, has-miR-548b-5p, has-miR-652-3p, and has-miR-193b-3p). All miRNAs which were predicted to be significantly dysregulated by the small RNA sequencing, except for hsa-miR-193b-3p, were confirmed to be dysregulated by RT-PCR in the Oslo cohort II cDNA, as shown in **Figure 19A**.

To unravel which are the hsa-miR-576-5p target genes, an available online tool, TargetScan, and two candidate genes targeted by has-miR-576-5p are periostin (*POSTN*) and insulin-growth-factor 1 (*IGF1*). The TargetScan results for these genes are shown in **Table 8**. Next, the expression of *POSTN* and *IGF1* at the transcript level was verified by RT-PCR in Oslo Cohort II and the Charité Cohort placentae samples, as shown in **Figure 19B** and **C**. In Oslo Cohort II samples, POSTN did not show a significant downregulation in EOPE or LOPE patient placentae versus the control, whereas IGF1 was significantly downregulated in EOPE versus control placentae, and not significantly downregulated in LOPE placentae versus control. However, in the Charité cohort which contains gestational age-matched controls, but much less samples, neither POSTN or IGF1 were downregulated significantly in EOPE placentae but was significantly downregulated in LOPE placentae (p-value < 0.05), and POSTN showed a downregulated tendency but insignificantly.

Table 12: Lists of top dysregulated (upregulated and downregulated) miRNA candidates from small RNA sequencing data in EOPE placentae from Oslo Cohort II. Legend: bolded = candidates for validation by RT-PCT, bolded and red = confirmed to be dysregulated by RT-PCR in this study, * = dysregulation well characterized in literature, ** = TE derived miRNA.

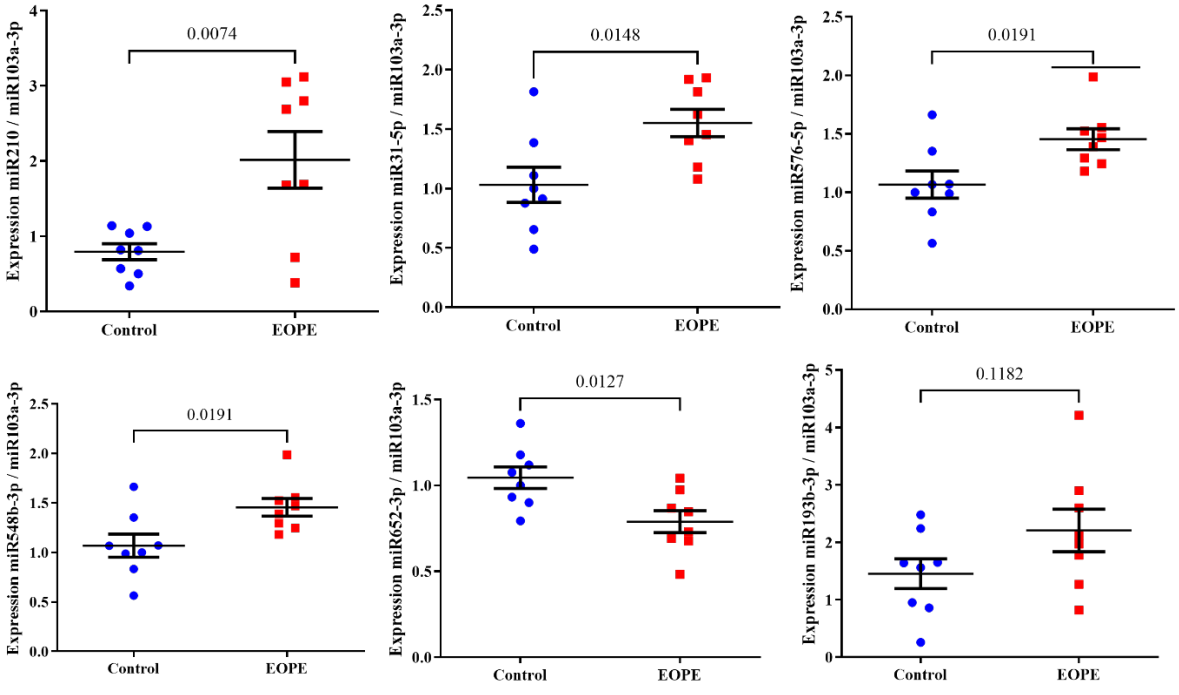
Upregulated miRNA in PE Placentae

Name	Fold Change	P-value
miR210*	1.423096165	8.61E-09
miR193b-3p	1.00376678	2.03E-05
miR31-5p	0.949219534	6.87E-05
miR548b-3p**	0.900765002	0.00036941
miR365a-3p	1.196299535	0.001381327
miR873-3p	1.111759872	0.002016335
miR873-5p	1.068422294	0.002167172
miR187-3p	1.343744621	0.002167172
miR876-5p	0.940172132	0.003266859
miR193b-5p	1.034884959	0.003699934
miR-576-5p**	0.450835681	0.014082415

Downregulated miRNA in PE Placentae

Name	Fold Change	P-value
miR194-5p	-0.507916627	2.03E-05
miR190b	-0.716408456	0.000272083
miR223-3p	-0.578806728	0.000524998
miR652-3p**	-0.521777186	0.00059438
miR221-5p	-0.447727505	0.002679553
miR1301	-0.629064423	0.002789638
miR-137	-2.001989672	0.004043251
miR-223-5p	-0.698285038	0.004790764
mir-146a-5p	-0.59437016	0.007337258

A)



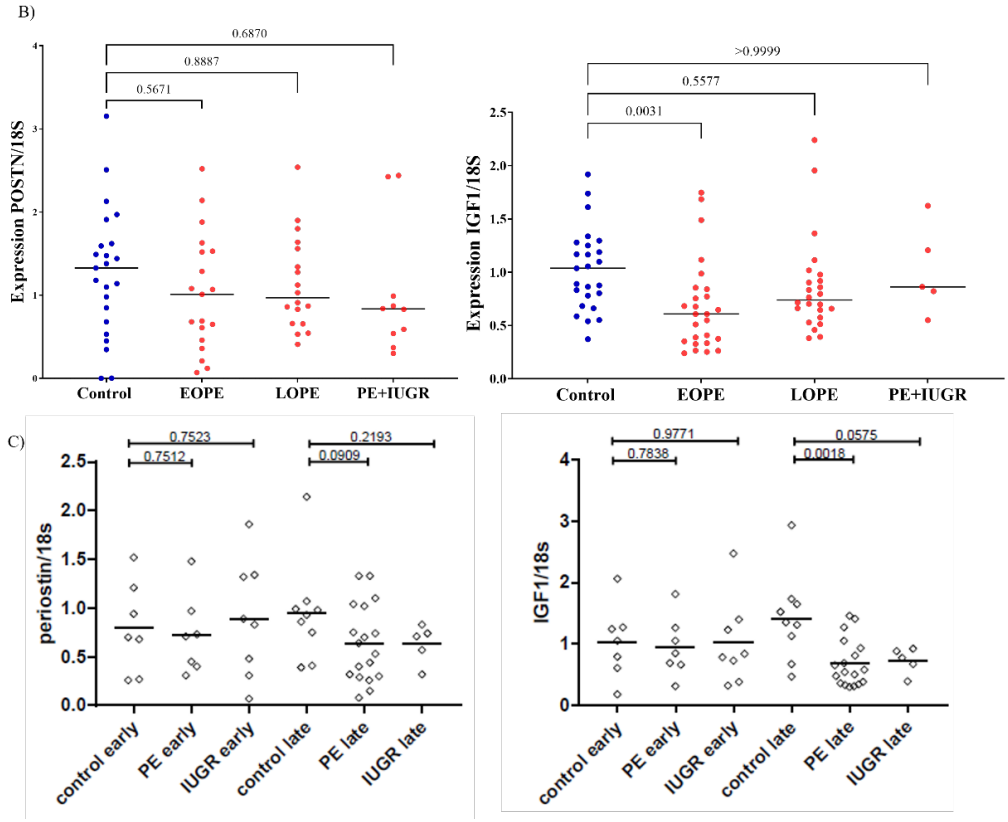


Figure 19: RT-PCR expression data of dysregulated miRNAs and the target genes of hsa-miR-576-5p in human placentae. A) RT-PCR data of selected six miRNAs probed in control (N=8) versus EOPE (N=8) placentae samples of Oslo Cohort II. The statistical calculation was done by an unpaired t-test. **B)** RT-PCR data of hsa-miR-576-5p candidate target genes, i: *POSTN* and ii: *IGF1*, in control (i: N=23; ii: N=25) versus EOPE (i. N=19; ii: N=26), LOPE, (i. N=18; ii: N=24) and PE+IUGR (i. N=10; ii: N=5) placentae samples of Oslo Cohort II. The statistics were calculated with Ordinary One-way ANOVA. **C)** POSTN and IGF1 were also probed in early or late pregnancy control versus early or late EOPE, LOPE, and PE+IUGR, respectively, in placenta samples of the Charité Cohort. The statistics were calculated with Ordinary One-way ANOVA.

4.6.1. The physiological role of POSTN and IGF1 in human trophoblast cells

To characterize a potential physiological impact of a dysregulated LINE-1 derived miRNA set of target genes, a transwell invasion assay (**Figure 20A**) was conducted using SGHPL4 cells and induced with the POSTN and IGF1, individually, or in combination, as chemoattractants. Based

on assay optimization, 100ng and 500ng of individual chemoattractants or combined in culture medium and 5% FBS were used for the assay, and the effects were compared to conditions containing only medium and 5% FBS. Upon 24h treatment, 500ng of POSTN only, and 500ng of POSTN and IGF1, had a significant effect (p-value <0.05) on SHGPL4 increase in invasion, while 500ng of IGF1 showed a tendency to affect invasion, but insignificantly, as shown in **Figure 20B**.

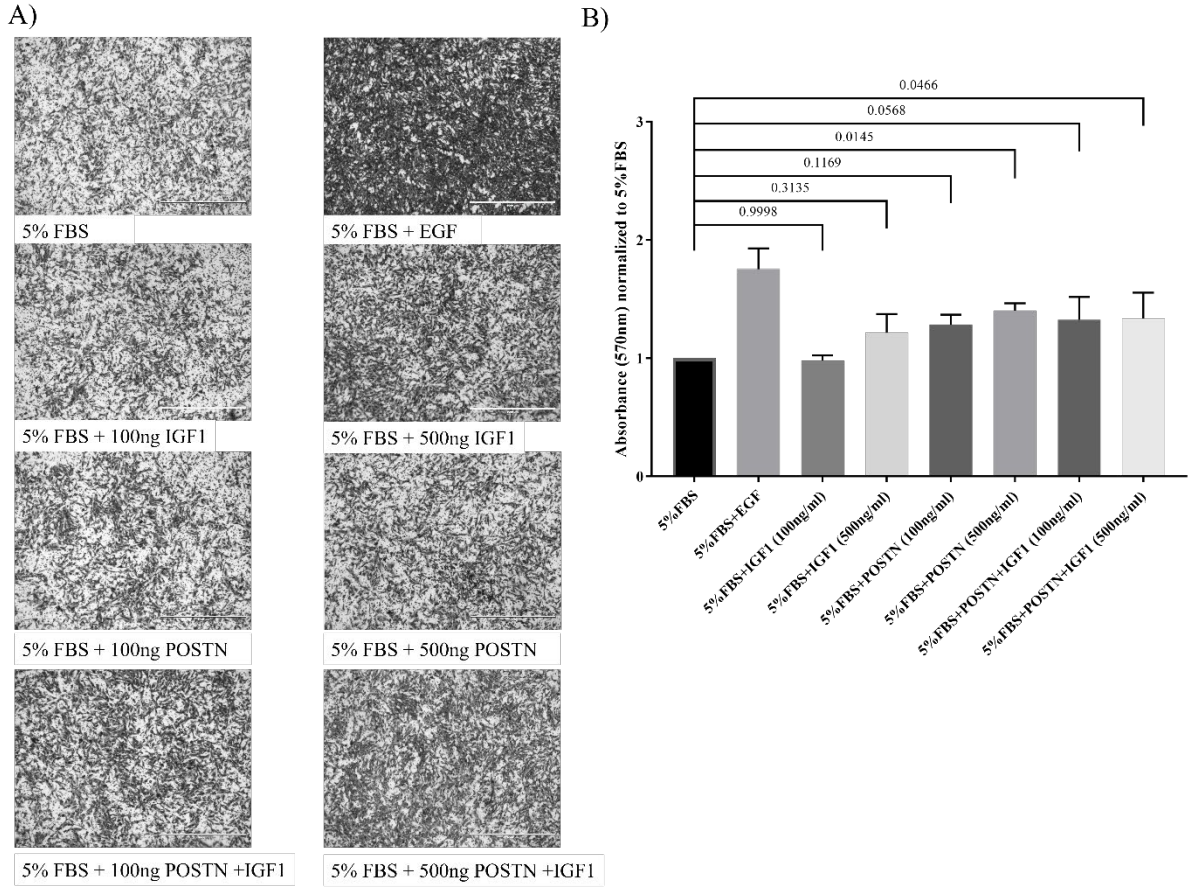


Figure 20: Invasion assay in SGHPL4 cells induced by POSTN and IGF1 as chemoattractants. A) An example from one experiment of transwell images of invaded cells after 24h subjection to chemoattractants. The images were acquired in bright field and 4X magnification. **B)** Colorimetric quantification of the transwell invasion assay. Each condition was compared to the 5% FBS control, and the experiments were repeated 3 times. The statistics were calculated using the Ordinary one-way ANOVA test.

5. DISCUSSION

5.1. LINE-1 expression and activity in the human placenta

The placenta is one of the ideal organs for LINE-1 retrotransposition. Firstly, in the placenta, the DNA methylation levels are decreased compared to any other somatic tissue, and the methylation level is comparable to a tumor environment [210], which is relatable to the >1000 publications discussing LINE-1 overexpression and activity in cancer [211]. Secondly, the placenta is a short-lived organ but a rapidly growing one, which therefore may tolerate a level of genome instability as a trade-off for resources saved on suppressing LINE-1 elements, such as DNA methylation, a very energy costly process [212]. This leads to the idea that the placenta fine tunes the tolerance of LINE-1 expression and activity, meanwhile benefitting from having a relaxed epigenetic landscape. Moreover, it is known that TEs played an important role in placenta evolution and development. The well-known example of TE-derived genes in placentation is the syncytins derived from retroviral envelope genes, which have co-opted the function for trophoblast fusion and the development of STBs [213]. Since somatic transposition events are not trans-generationally inherited, this activity holds no apparent benefit to the TE. Is is speculated that regulated somatic TE activity may reflect a symbiotic relationship between host and TE [8]. Marcia et al. have suggested that LINE-1 activity in somatic cells does not represent a clear selective advantage for the TEs, but the created intra-individual heterogeneity contributes to the plasticity and fitness of the host [37].

This research characterized for the first time the expression of LINE-1 in the human placenta. I showed that the evolutionary youngest LINE-1 elements, L1HS, are expressed on the RNA level by RNA sequencing and RT-PCT, as early as the TE-lineage in the pre-implantation embryo and placental CTBs (**Figures 12** and **13**). On the protein level, all three subtypes of trophoblast cells of the human placenta expressed L1ORF1 while it was not expressed in the other cell types of the placenta, as also seen in the pregnancy-related tissue panel RT-PCR (**Figure 12B**) and IHC staining of the placenta (**Figure 15E** and **F**). It was also observed, but not quantified, that L1ORF1 is more expressed in the first-trimester placenta samples versus the third trimester (**Figure 15F**). The different forms of L1ORF1 protein detection in Western Blots as shown in **Figure 15A-C** were similarly previously reported by Sokolowski et al. Who have also identified

multiple expressed forms of the OFR1 protein, that is multiple truncated versions of the proteins, monomers, and multiple dimer and trimer combinations [214]. The comparison of L1ORF1 protein trimer (the retrotranspositionally active form) levels did not reveal any significant difference in expression (**Figure 15D**), suggesting that either L1ORF1 is not dysregulated at the protein level or that more likely the selected placenta samples were not from the subtype that expressed high L1HS. In SGHPL4 EVTs, subcellular localization of L1ORF1 was detected in the cytoplasm and the nucleus. As observed in the IF staining (**Figure 15G**), Goodier et al. have also reported that L1ORF1 forms large cytoplasmic foci, which they identified to be stress granules. However, it remains to confirm that the punctate sequestering of L1ORF1 are stress granules in humans EVTs. Recent findings also showed that distinct nuclear and cytoplasmic L1-RNP complexes exist, but with ORF1p-rich RNPs predominantly found in the cytoplasm [215, 216], which is per the findings of this study. It is also speculated that even a few L1ORF1 localization in the nucleus may be a mark of active retrotransposition [209], as also observed in the IF results of this research (**Figure 15G**).

LINE-1s are the only known autonomous retrotransposons to be active in the human genome. determined. As shown in **Figure 16**, LINE-1 activity via a retrotransposition assay was only confirmed *in vitro* in the EVT cell line, SGHPL4, but not in primary human samples. The retrotransposition rate in SGHPL4 and DLX5OE-SGHPL4 cells was low; however, this is an expected rate as it was reported by Ostertag et al. that L1RP-EGFP retrotransposes in the assay at a rate of about 0.5% of transfected cells/ day, while no retrotransposition events are detected within the first 48 hours of the assay [217]. As previously mentioned, for retrotransposition to occur, both ORF proteins are required to be expressed [51]. In the Western blot analysis of this study, only L1ORF1 was confidently detected. Due to the poor quality of the few existing L1ORF2 antibodies on the market and the very low expression of the endogenous L1ORF2 protein, L1ORF2 could not be confidently detected in human primary samples, as also reported by others [218]. Recently, Mukherjee et al have reported to have developed a very specific polyclonal anti-L1ORF2 antibody, which specifically detects L1ORF2 in human post-operative oral tissue.

To ultimately confirm the activity of LINE-1 in human trophoblasts, retrotransposon capture (RC) sequencing should be employed [219]. In this method, the fetal trophoblast genome is compared to the somatic sample genome from the same individual, and it is also compared to the parental somatic genome. This way, new placenta-specific LINE-1 integrations are normalized to

the pre-existing LINE-1 retrotransposition events in the parental genomes and fetal early embryogenesis, indicating definite LINE-1 retrotransposition activity.

5.2. LINE-1s and human placenta specificity

Since the human-specific LINE-1 elements are only dysregulated in a subset of third trimester PE patients, as shown in **Figure 13B**, these results are insufficient to support the hypothesis that dysregulated human-specific TEs are the cause of human-specificity of PE. However, the hypothesis also cannot be entirely rejected as LINE-1 elements are more highly expressed in the first-trimester placentae; therefore their role in the first trimester of PE pathogenesis should be explored further. Considering that developing first-trimester placenta samples are inaccessible, perhaps a secreted readout of LINE-1 expression, such as LINE-1-derived miRNAs, or activity would need to be tested. Nevertheless, since the human-specific L1Hs are generally active in the placenta, this may support the idea that placental TE activity has contributed to the human-specificity evolution of the human versus ancestral placentas. Recent few studies on the evolution of the placenta agree that the last common ancestor of primates had an invasive, hemochorial placenta; therefore, the placenta is anatomically conserved with our recent common ancestors [220-222]. However, the placenta is a fast-evolving organ, meaning that it is one of the most variable organs within mammals. Hou et al. identified 94 highly expressed genes in the human placenta that have evolved adaptively during human evolution since the time of the last common ancestor of eutherian mammals [223]. Interestingly, immune response genes at the maternal-fetal interface, such as killer-cell immunoglobulin-like receptors (KIRs), as previously discussed to be very important in implantation, have also been subject to intense selective pressures [224]. As shown in this study that LINE-1 element overexpression impacts immune response genes in the placenta, it could be speculated that the general human-specific LINE-1 expression and activity in the human placenta has pushed forward the evolution of the immune response at the maternal-fetal interface.

5.3. Regulation of LINE-1 in the placenta and PE

Many different genetic elements and molecular pathways are dysregulated in different PE samples; therefore it can be said that PE is a heterogeneous disorder and is most likely varying in molecular origin. Due to the heterogeneity of PE, it is of no surprise that only a subgroup of the PE patients displays L1HS overexpression, as indicated by the RNA sequencing results (**Figure 13B**). It would be important further to characterize the LINE-1 regulatory mechanisms in this subgroup. When analyzing the published list of LINE-1 regulators [96] in the total PE patient transcriptome (**Figure 18**), 11 genes are dysregulated, however, according to the casTLE effect score of the integrated study [96], these genes do not have a very high regulation effect score. This may be since LINE-1s are not generally dysregulated across PE patients but only in a subgroup. **Table 11** shows the analysis of the same set of regulators in the LINE-1-high expressing group (3 PE versus 3 Control), other regulators, such as SAMHD1 and RAB36, that have stronger casTLE scores are significantly dysregulated in LINE-1 high PE versus LINE-1 high controls patient trophoblasts. SAMHD1 enhances sequestration of LINE-1 RNP complex in stress granules and consequent attenuation of LINE-1 retrotransposition [225]. The role of the small GTPase, RAB36, in LINE-1 regulation has not been reported on so far. Moreover, these two regulators are interesting candidates to be focused on in further studies of this PE subtype group. Since the RNA sequencing was conducted on third-trimester patient placenta samples, it may be the case that the upregulated repressor genes (SAMHD1, OPRK1, and EZH2) are a result of a negative feedback loop, as an attempt to attenuate LINE-1 expression. In the retrotransposition assay, quantification of GFP positive cells, indicating one round of successful retrotransposition [226], showed that DLX5OE-SGHPL4 cells (PE model cell line) have higher LINE-1 activity. Based on expression correlation of L1HS and DLX5 in the Oslo Cohort II data, their expression correlates significantly. This result raises the question if LINE-1 is upregulated on the transcriptional level upon DLX5 overexpression, as DLX5 is a transcription factor, or does DLX5 affect LINE-1 activity indirectly. Zadora et al. have published microarray data of DLX5OE-SGHPL4 cells [127], however the expression of TEs cannot determined from microarray data, and therefore effect of DLX5 on L1HS expression remains to be clarified.

5.4. An immune-mediated PE subtype

Pregnancy is described as a mild inflammation [227]. The fetal-maternal interface is challenged to maintain the delicate balance between cell proliferation and invasion for proper placentation and fetal development versus uncontrolled invasion, which is in part controlled by inflammation. If the maternal immune system is over-stimulated, the exaggerated immune response restricts the extent of invasion and remodeling of maternal spiral arteries, leading to poor vascularization of the placenta. Therefore, it is suggested that a mild inflammatory response is necessary for a successful pregnancy [228], which supports the hypothesis that this may be a co-opted function of controlled LINE-1 expression and activity in the placental trophoblasts.

Cytokines, which are produced and secreted by trophoblast cells, play an essential role in signaling between immune cells, and they can be categorized as pro- (such as IL-2, TNFα, and INFγ) or anti-inflammatory (GM-CSF, IL-3, and IL-10). Elevated IL-12 and IL-18 have been observed in PE placentae, while IL-10 is a decrease in the anti-inflammatory cytokines [229, 230]. It has been proposed that cytokines produced by the Th1-type immune reaction might attack the fetal (semi-allograft) trophoblasts [229]. In PE, it is believed that IL-18 and IL-12 in combination together change the immune reactivity equilibrium towards the Th1 phenotype [230].

A study by Benton et al. assessed histopathological lesions in the placentae of five different gene expression-based PE clusters [231]. They identified that one of the five PE types, and most heterogeneous in gene expression, was the chronic inflammation subtype, demonstrating overexpression of TNFα, IFNγ, CXCL10. Correlation data of this research showed that L1HS expression does not correlate with the current prognostic marker, the sFlt/PlGF ratio, at least not in the third trimester, suggesting that not all PE subtypes can be detected using the current marker. In the future, these patients would potentially benefit from having LINE-1 or inflammation-related secreted biomarkers investigated in the first trimester as a possible early diagnostic tool or treatment of the disorder.

Moreover, many recent studies have reported on the impact of LINE-1 on DNA damage, cell growth, cell senescence, and type I interferon (INF) pathways [111, 232-234]. LINE-1 retrotransposition creates DNA double-stranded breaks, stimulating DNA damage response genes, activating cell-cycle arrest via ATM expression, and inducing the INF response [232, 235].

Ardeljan et al. showed that adding LINE-1 reverse-transcriptase inhibitors, such as zalcitabine or didanosine, attenuated the IFN response [111]. That is per the finding of this research, where PE with high LINE-1 expression in their trophoblasts also exhibited an upregulation in many DNA damage, type I interferon, and cell senescence pathway genes. Similar was reported in a single cell sequencing study of three PE and three control placentae, revealing that CTBs and EVTs in PE placentae exhibit mainly an upregulation in genes involved in the immune response [236]. Other than exaggerated immune cell response and inflammation, increased cellular senescence in EVTs may contribute to PE by preventing cell proliferation and therefore decreased organ growth [111]. A mechanism for decreased cell proliferation in human EVTs expressing high levels of DLX5 has been reported by Zadora et al., who also showed that DLX5 is upregulated in 70% of PE patients [127].

5.5. The indirect impact of LINE-1 on physiology during pregnancy

Dysregulated TEs differ in behavior and impact versus the rest of genomic entities since even one dysregulated TE loci can have a downstream impact on the genome stability and local gene expression. For example, in the case of TE-derived miRNA has-miR-576-5p, it is derived from a LINE-1 locus. Suppose the expression of that LINE-1 locus is upregulated. In that case, the embedded miRNA is also upregulated, or if that LINE-1 locus retrotransposes, it also amplifies the copy number of that miRNA sequence. In turn, these events have an impact on the downstream miRNA-regulated genes. It should be noted that this impact is further amplified since miRNAs are secreted molecules that affect the interacting surrounding tissue and blood circulation. Such a scenario was tested in this research as represented by **Figures 19 and 20**. The LINE-1 element that carries has-miR-576-5p is L1MB7, a highly conserved LINE-1 element among mammals [237]. Since it is a very ancient LINE-1 element, it contains acquired mutations and does not produce a full-length transcript in humans, which is necessary for LINE-1 activity, however, its regulation may have an effect on other genetic entities that it may harbor, such as miR-576-5p. The effect of hsa-miR-576-5p on POSTN and IGF1 remains to be elucidated; however since the target genes are not dysregulated in EOPE while has-miR-576-5p is, this suggests that even if the miRNA has an impact on the expression of these genes, then it is not a strong effect on its own, rather requires a synergistic effect with other miRNAs. Regardless, it was important to explore the potential involvement of POSTN and IGF1 in PE pathogenesis because these genes are involved

in pathways that are highly important for EVT function, mainly in αvβ3 integrin-mediated cell adhesion and mobility [238-240]. Previous studies on POSTN have revealed that POSTN is progesterone regulated, impacts migration and attachment of ovine trophectoderm, and is an important mediator of implantation [241]. They are both also secreted molecules into the maternal circulation and surrounding tissue; therefore, they may not only have biomarker potential but also could signal and impact the surrounding tissue environment. It was reported that POSTN could potentially be a biomarker of miscarriage when higher levels are detected in the fourth week of pregnancy versus the sixth week in serum samples of pregnant women [242]. According to the data of this study (**Figure 20C**), POSTN and IGF1 may be interesting candidates to investigate in LOPE pathogenesis, as their expression is decreased at the transcription level in PE placentae versus the age-matched control samples.

5.6. Limitation of PE research

PE, as it is a disease that primarily involves the placenta, a transient organ, would seem to be an ideal system to study the molecular mechanism of a disease, since the tissue is by default available. However, the remaining main issue of studying the origin of PE is that the disease initiates very early in pregnancy with the first symptoms detected only from the 20th week of gestation, and during this time, the tissue cannot be sampled. So far, the only available early pregnancy placenta or decidua tissue for studying PE molecular or genetic mechanisms is sampled from voluntary abortions. This is not ideal since it is not known at that stage whether there is a risk in this patient to develop any pregnancy-related diseases such as gestational diabetes, hypertension, IUGR, or PE. Other available and related tissues are maternal blood and urine samples; however, they only contain secreted biomarkers of the diseases in the mother. Whole placenta or decidua samples retrieved after PE-related emergency C-section, are of great value and have provided much information about the nature of the disease, however at this stage in pregnancy such samples could not be probed for questions regarding the disease origin and development.

A complete pregnancy sample cohort is an unmet need. This includes samples from the mother, father, and the baby, and samples such as placenta, decidua, blood (sample per trimester), maternal urine (per trimester), as well as the maternal clinical parameters throughout pregnancy and medical history, fetal blood or skin sample (reference tissue), fetal clinical parameters *in utero* and after

birth, father's sperm sample, blood, and medical history. Such samples would allow for comparative genetics, epigenetics, and biochemistry between the parents and the offspring in healthy versus problematic pregnancies. However, due to the inaccessibility of the placenta or fetus during development *in utero*, catching the disease before PE onset will remain an obstacle unless early biomarkers are discovered in tissue that can be sampled during pregnancy for all of the subtypes of PE.

Another limitation in PE research is that it is so far determined to be only a human-specific disease. Animal models are essential to identify the disease origins, develop new therapies, and conduct pre-clinical evaluations for the safety and efficacy of interventions such as long-term effects on the mother's and baby's health. Different approaches across multiple species have been used to generate preclinical models of PE, where maternal hypertension, proteinuria, renal pathology, impaired remodeling of blood vessels supplying the uterus, placental dysfunction, and/or IUGR are present. The most well-known PE animal models are the reduced uterine perfusion pressure (RUPP) PE model in rats and the genetically modified models of PE in mice [243, 244]. A review by Gatford et al. summarizes that all PE animal models so far differ from the human condition and therefore provide limited information since experimental PE is induced by surgical, pharmacological, immunological, or genetic manipulation before or during pregnancy. In contrast, in humans, PE occurs spontaneously in only some women and with differing sub-phenotypes [245]. The main limitation in these models is to study underlaying the human-specific genetics and epigenetics that predisposes the women, or particular pregnancies, to PE. It would be relevant to exogenously express L1ORFeus in early rat embryos and test whether L1HS expression impacts the rat placenta towards more human-like features, particularly on the immunological level.

5.7. Summary and Outlook

I have successfully characterized the expression profile of LINE-1 elements in the placenta, on the RNA and protein levels, as well as on the tissue, cell, and sub-cellular levels.My data suggests that LINE-1 element expression and activity may be tolerated in the placenta since it is a short-lived organ without any mutation inheritance risk, meanwhile the presentce of a certain dose of LINE-1 expression and activity may have been befeticial for the placenta towards generating a mild-inflammatory state. I further provided data towards the elucidation of the immune-mediated

PE subtype with the involvement of LINE-1 expression and over-expression in immune regulation during pregnancy and PE. I showed that the SGH-PL4 EVT cell line can support LINE-1 activity and that LINE-1s are significantly more active in the PE model cell line versus wildtype, suggesting that LINE-1 are not just expressed in the placenta but also retrotranspositionally active. I provided insight into possible LINE-1 regulator candidates in the placenta and PE. Lastly, I have identified previously unreported dysregulated TE-derived and other miRNAs in PE vs. control placenta samples.

For the future continuation of this project, it would be important to understand better the molecular impact of the transcription factor DLX5 on LINE-1 expression or activity in placenta development and PE. The remaining question is does DLX5 overexpression cause directly or indirectly an upregulation of L1HS, or does it only impact the cellular environment that promotes LINE-1 retrotransposition. With the knowledge of the DLX5 target sequence, it would be important to analyze all active LINE-1 loci to determine if they contain a DLX5 binding sequence and test one of those loci *in vitro* with the DLX5 sequence knockout whether the expression of LINE-1 is affected.

In LINE-1 research, the RC-seq tool is an essential tool for determining LINE-1 activity. It would be important to conduct RC-seq in PE versus healthy trophoblast samples with the above-mentioned appropriate controls to detect novel placenta-specific LINE-1 integration sites and quantify them. Additionally, ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) on the same trophoblast genomic DNA as the RC-seq analysis should be conducted to understand the epigenetic landscape in those individuals detect the open chromatin regions in the genome. Ideally, such experiments would be conducted on samples that were previously analyzed by RNA sequencing in which LINE-1 high expressing patients were determined. These three experiments in the same set of samples would provide an invaluable set of data for further understanding LINE-1 biology and PE pathogenesis.

In PE research, the immune-mediated PE subtype model should be explored further. Upon LINE-1 induction, the amount of DNA damage should be quantified and checked if it is proportional to the severity of the immune response expressing genes and cell senescence. Also the effect of blocking LINE-1 retrotransposition by pharmaceutically available RT-inhibitos would give insight into the level of effect that LINE-1 has on initiating the innate immune response gene expression and cell senescence, as well as impacting EVT functions such as proliferation and

invasion. The profile of secreted cytokines should also be examined to elucidate a potential early biomarker readout of this disorder. Interestingly, an interesting candidate an intervention in this PE subtype would be by probiotic treatment, where various underlying mechanisms for immunomodulation have been suggested, including altering the Th1/Th2 balance and T-reg cell production [246]. Further studies on the molecular level should be conducted to well characterize this subtype towards identifying possible medical treatment.

6. REFERENCES

1. Orgel, L.E. and F.H. Crick, *Selfish DNA: the ultimate parasite.* Nature, 1980. **284**(5757): p. 604-7.
2. Doolittle, W.F. and C. Sapienza, *Selfish genes, the phenotype paradigm and genome evolution.* Nature, 1980. **284**(5757): p. 601-3.
3. Kazazian, H.H., Jr., et al., *Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man.* Nature, 1988. **332**(6160): p. 164-6.
4. Cordaux, R. and M.A. Batzer, *The impact of retrotransposons on human genome evolution.* Nat Rev Genet, 2009. **10**(10): p. 691-703.
5. Mc, C.B., *The origin and behavior of mutable loci in maize.* Proc Natl Acad Sci U S A, 1950. **36**(6): p. 344-55.
6. Britten, R.J. and D.E. Kohne, *Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms.* Science, 1968. **161**(3841): p. 529-40.
7. Britten, R.J. and E.H. Davidson, *Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty.* Q Rev Biol, 1971. **46**(2): p. 111-38.
8. Chuong, E.B., N.C. Elde, and C. Feschotte, *Regulatory activities of transposable elements: from conflicts to benefits.* Nat Rev Genet, 2017. **18**(2): p. 71-86.
9. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
10. de Koning, A.P., et al., *Repetitive elements may comprise over two-thirds of the human genome.* PLoS Genet, 2011. **7**(12): p. e1002384.
11. Anthony J. F. Griffiths, S.R.W., Richard C. Lewontin, Sean B. Carroll, *Introduction to Genetic Analysis, 9th Edition*. 9 ed.: W. H. Freeman and Company.
12. Cosby, R.L., N.C. Chang, and C. Feschotte, *Host-transposon interactions: conflict, cooperation, and cooption.* Genes Dev, 2019. **33**(17-18): p. 1098-1116.
13. Gilbert, C. and C. Feschotte, *Horizontal acquisition of transposable elements and viral sequences: patterns and consequences.* Curr Opin Genet Dev, 2018. **49**: p. 15-24.
14. Van Valen, L., *A new evolutionary law.* Evolutionary Theory, 1973. **1**: p. 1-30.
15. Lisch, D. and J.L. Bennetzen, *Transposable element origins of epigenetic gene regulation.* Curr Opin Plant Biol, 2011. **14**(2): p. 156-61.
16. Tristan-Ramos, P., et al., *The tumor suppressor microRNA let-7 inhibits human LINE-1 retrotransposition.* Nat Commun, 2020. **11**(1): p. 5712.
17. Yoder, J.A., C.P. Walsh, and T.H. Bestor, *Cytosine methylation and the ecology of intragenomic parasites.* Trends Genet, 1997. **13**(8): p. 335-40.
18. Huang, C.R., K.H. Burns, and J.D. Boeke, *Active transposition in genomes.* Annu Rev Genet, 2012. **46**: p. 651-75.
19. Bourque, G., et al., *Ten things you should know about transposable elements.* Genome Biol, 2018. **19**(1): p. 199.

20. Platt, R.N., 2nd, M.W. Vandewege, and D.A. Ray, *Mammalian transposable elements and their impacts on genome evolution.* Chromosome Res, 2018. **26**(1-2): p. 25-43.

21. Boeke, J.D., et al., *Ty elements transpose through an RNA intermediate.* Cell, 1985. **40**(3): p. 491-500.

22. Craig, N.L., Craigie, R., Gellert, M., Lambowitz, A. M., *Mobile DNA II*. 2002, Washington DC: ASM Press.

23. Kapitonov, V.V. and J. Jurka, *Helitrons on a roll: eukaryotic rolling-circle transposons.* Trends Genet, 2007. **23**(10): p. 521-9.

24. Pace, J.K., 2nd and C. Feschotte, *The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage.* Genome Res, 2007. **17**(4): p. 422-32.

25. Lower, R., J. Lower, and R. Kurth, *The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences.* Proc Natl Acad Sci U S A, 1996. **93**(11): p. 5177-84.

26. Boeke, J.D. and J.P. Stoye, *Retrotransposons, endogenous retroviruses, and the evolution of retroelements*. Retroviruses ed. 1997, Cold Spring Habor, New York: Cold Spring Harbor Laboratory Press.

27. Mills, R.E., et al., *Which transposable elements are active in the human genome?* Trends in genetics, 2007. **23**(4): p. 183-191.

28. Dewannieux, M., C. Esnault, and T. Heidmann, *LINE-mediated retrotransposition of marked Alu sequences.* Nat Genet, 2003. **35**(1): p. 41-8.

29. Boissinot, S., P. Chevret, and A.V. Furano, *L1 (LINE-1) retrotransposon evolution and amplification in recent human history.* Mol Biol Evol, 2000. **17**(6): p. 915-28.

30. Myers, J.S., et al., *A comprehensive analysis of recently integrated human Ta L1 elements.* Am J Hum Genet, 2002. **71**(2): p. 312-26.

31. Sheen, F.M., et al., *Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition.* Genome Res, 2000. **10**(10): p. 1496-508.

32. Batzer, M.A. and P.L. Deininger, *Alu repeats and human genomic diversity.* Nat Rev Genet, 2002. **3**(5): p. 370-9.

33. Wang, J., et al., *dbRIP: a highly integrated database of retrotransposon insertion polymorphisms in humans.* Human mutation, 2006. **27**(4): p. 323-329.

34. Xiong, Y. and T.H. Eickbush, *Origin and evolution of retroelements based upon their reverse transcriptase sequences.* EMBO J, 1990. **9**(10): p. 3353-62.

35. Smit, A.F., et al., *Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences.* J Mol Biol, 1995. **246**(3): p. 401-417.

36. Furano, A.V., *The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons.* Prog Nucleic Acid Res Mol Biol, 2000. **64**: p. 255-94.

37. Macia, A., et al., *Activity of retrotransposons in stem cells and differentiated cells*, in *Human Retrotransposons in Health and Disease*. 2017, Springer. p. 127-156.

38. Grimaldi, G., J. Skowronski, and M.F. Singer, *Defining the beginning and end of KpnI family segments.* EMBO J, 1984. **3**(8): p. 1753-9.

39. Szak, S.T., et al., *Molecular archeology of L1 insertions in the human genome.* Genome Biol, 2002. **3**(10): p. research0052.

40. Doucet-O'Hare, T.T. and H.H. Kazazian, *Retrotransposon Contribution to Genomic Plasticity*, in *Human Retrotransposons in Health and Disease*. 2017, Springer. p. 63-93.

41. Brouha, B., et al., *Hot L1s account for the bulk of retrotransposition in the human population.* Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5280-5.

42. Beck, C.R., et al., *LINE-1 retrotransposition activity in human genomes.* Cell, 2010. **141**(7): p. 1159-70.

43. Tuzun, E., et al., *Fine-scale structural variation of the human genome.* Nat Genet, 2005. **37**(7): p. 727-32.

44. Korbel, J.O., et al., *Paired-end mapping reveals extensive structural variation in the human genome.* Science, 2007. **318**(5849): p. 420-6.

45. Kidd, J.M., et al., *Mapping and sequencing of structural variation from eight human genomes.* Nature, 2008. **453**(7191): p. 56-64.

46. Xing, J., et al., *Mobile elements create structural variation: analysis of a complete human genome.* Genome Res, 2009. **19**(9): p. 1516-26.

47. Babushok, D.V. and H.H. Kazazian, Jr., *Progress in understanding the biology of the human mutagen LINE-1.* Hum Mutat, 2007. **28**(6): p. 527-39.

48. Scott, A.F., et al., *Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence.* Genomics, 1987. **1**(2): p. 113-25.

49. Swergold, G.D., *Identification, characterization, and cell specificity of a human LINE-1 promoter.* Mol Cell Biol, 1990. **10**(12): p. 6718-29.

50. Speek, M., *Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes.* Mol Cell Biol, 2001. **21**(6): p. 1973-85.

51. Moran, J.V., et al., *High frequency retrotransposition in cultured mammalian cells.* Cell, 1996. **87**(5): p. 917-27.

52. Denli, A.M., et al., *Primate-specific ORF0 contributes to retrotransposon-mediated diversity.* Cell, 2015. **163**(3): p. 583-93.

53. Holmes, S.E., M.F. Singer, and G.D. Swergold, *Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element.* J Biol Chem, 1992. **267**(28): p. 19765-8.

54. Martin, S.L., *Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells.* Mol Cell Biol, 1991. **11**(9): p. 4804-7.

55. Martin, S.L. and F.D. Bushman, *Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon.* Mol Cell Biol, 2001. **21**(2): p. 467-75.

56. Feng, Q., et al., *Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition.* Cell, 1996. **87**(5): p. 905-16.

57. Mathias, S.L., et al., *Reverse transcriptase encoded by a human transposable element.* Science, 1991. **254**(5039): p. 1808-10.

58. Fanning, T. and M. Singer, *The LINE-1 DNA sequences in four mammalian orders predict proteins that conserve homologies to retrovirus proteins.* Nucleic Acids Res, 1987. **15**(5): p. 2251-60.

59. Criscione, S.W., et al., *Genome-wide characterization of human L1 antisense promoter-driven transcripts.* BMC Genomics, 2016. **17**: p. 463.

60. Roman-Gomez, J., et al., *Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia.* Oncogene, 2005. **24**(48): p. 7213-23.

61. Deininger, P.L., et al., *Master genes in mammalian repetitive DNA amplification.* Trends Genet, 1992. **8**(9): p. 307-11.

62. Wei, W., et al., *Human L1 retrotransposition: cis preference versus trans complementation.* Mol Cell Biol, 2001. **21**(4): p. 1429-39.

63. Kulpa, D.A. and J.V. Moran, *Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition.* Hum Mol Genet, 2005. **14**(21): p. 3237-48.

64. Hayashi, Y., et al., *Mechanism by which a LINE protein recognizes its 3′ tail RNA.* Nucleic Acids Res, 2014. **42**(16): p. 10605-17.

65. Monot, C., et al., *The specificity and flexibility of l1 reverse transcription priming at imperfect T-tracts.* PLoS Genet, 2013. **9**(5): p. e1003499.

66. Alisch, R.S., et al., *Unconventional translation of mammalian LINE-1 retrotransposons.* Genes Dev, 2006. **20**(2): p. 210-24.

67. Smit, A.F., *Interspersed repeats and other mementos of transposable elements in mammalian genomes.* Curr Opin Genet Dev, 1999. **9**(6): p. 657-63.

68. Kubo, S., et al., *L1 retrotransposition in nondividing and primary human somatic cells.* Proc Natl Acad Sci U S A, 2006. **103**(21): p. 8036-41.

69. Luan, D.D., et al., *Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition.* Cell, 1993. **72**(4): p. 595-605.

70. Kojima, K.K., *Different integration site structures between L1 protein-mediated retrotransposition in cis and retrotransposition in trans.* Mob DNA, 2010. **1**(1): p. 17.

71. Grimaldi, G. and M.F. Singer, *Members of the KpnI family of long interspersed repeated sequences join and interrupt alpha-satellite in the monkey genome.* Nucleic Acids Res, 1983. **11**(2): p. 321-38.

72. Becker, K.G., et al., *Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element.* Hum Mol Genet, 1993. **2**(10): p. 1697-702.

73. Coufal, N.G., et al., *L1 retrotransposition in human neural progenitor cells.* Nature, 2009. **460**(7259): p. 1127-1131.

74. Yang, N., et al., *An important role for RUNX3 in human L1 transcription and retrotransposition.* Nucleic Acids Res, 2003. **31**(16): p. 4929-40.

75. Tiwari, B., et al., *p53 directly represses human LINE1 transposons.* Genes Dev, 2020.

76. Sun, X., et al., *Transcription factor profiling reveals molecular choreography and key regulators of human retrotransposon expression.* Proc Natl Acad Sci U S A, 2018. **115**(24): p. E5526-E5535.

77. Goodier, J.L., *Restricting retrotransposons: a review.* Mob DNA, 2016. **7**: p. 16.

78. Slotkin, R.K. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome.* Nat Rev Genet, 2007. **8**(4): p. 272-85.

79. Bulut-Karslioglu, A., et al., *Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells.* Mol Cell, 2014. **55**(2): p. 277-90.

80. Jonsson, M.E., et al., *Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors.* Nat Commun, 2019. **10**(1): p. 3182.

81. Thayer, R.E., M.F. Singer, and T.G. Fanning, *Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1-encoded protein.* Gene, 1993. **133**(2): p. 273-7.

82. Min, B., et al., *Dnmt1 binds and represses genomic retroelements via DNA methylation in mouse early embryos.* Nucleic Acids Res, 2020. **48**(15): p. 8431-8444.

83. Kato, Y., et al., *Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse.* Hum Mol Genet, 2007. **16**(19): p. 2272-80.

84. Bourc'his, D. and T.H. Bestor, *Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L.* Nature, 2004. **431**(7004): p. 96-9.

85. Yu, F., et al., *Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription.* Nucleic Acids Res, 2001. **29**(21): p. 4493-501.

86. Fan, H., et al., *BAHCC1 binds H3K27me3 via a conserved BAH module to mediate gene silencing and oncogenesis.* Nat Genet, 2020. **52**(12): p. 1384-1396.

87. Ren, W., et al., *DNMT1 reads heterochromatic H4K20me3 to reinforce LINE-1 DNA methylation.* Nat Commun, 2021. **12**(1): p. 2490.

88. Anant, S. and N.O. Davidson, *Molecular mechanisms of apolipoprotein B mRNA editing.* Curr Opin Lipidol, 2001. **12**(2): p. 159-65.

89. Muckenfuss, H., et al., *APOBEC3 proteins inhibit human LINE-1 retrotransposition.* J Biol Chem, 2006. **281**(31): p. 22161-22172.

90. Hamdorf, M., et al., *miR-128 represses L1 retrotransposition by binding directly to L1 RNA.* Nat Struct Mol Biol, 2015. **22**(10): p. 824-31.

91. Yang, N. and H.H. Kazazian, Jr., *L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells.* Nat Struct Mol Biol, 2006. **13**(9): p. 763-71.

92. Cook, P.R., C.E. Jones, and A.V. Furano, *Phosphorylation of ORF1p is required for L1 retrotransposition.* Proc Natl Acad Sci U S A, 2015. **112**(14): p. 4298-303.

93. Castro-Diaz, N., et al., *Evolutionally dynamic L1 regulation in embryonic stem cells.* Genes Dev, 2014. **28**(13): p. 1397-409.

94. Wolf, G., D. Greenberg, and T.S. Macfarlan, *Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Kruppel-associated box zinc finger protein family.* Mob DNA, 2015. **6**: p. 17.

95. Robbez-Masson, L., et al., *The HUSH complex cooperates with TRIM28 to repress young retrotransposons and new genes.* Genome Res, 2018. **28**(6): p. 836-845.

96. Liu, N., et al., *Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators.* Nature, 2018. **553**(7687): p. 228-232.

97. Fukuda, K., et al., *A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells.* Genome Res, 2018. **28**(6): p. 846-858.

98. Tchasovnikarova, I.A., et al., *GENE SILENCING. Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells.* Science, 2015. **348**(6242): p. 1481-1485.

99. Tchasovnikarova, I.A., et al., *Hyperactivation of HUSH complex function by Charcot-Marie-Tooth disease mutation in MORC2.* Nat Genet, 2017. **49**(7): p. 1035-1044.

100. Lee, E., et al., *Landscape of somatic retrotransposition in human cancers.* Science, 2012. **337**(6097): p. 967-71.

101. Kondo-Iida, E., et al., *Novel mutations and genotype-phenotype relationships in 107 families with Fukuyama-type congenital muscular dystrophy (FCMD).* Hum Mol Genet, 1999. **8**(12): p. 2303-9.

102. Divoky, V., et al., *Unstable Santa Ana hemoglobin or alpha 2 beta 2 88 (F4) Leu-Pro detected in a Slovak girl.* Vnitr Lek, 1996. **42**(4): p. 258-61.

103. Schwahn, U., et al., *Positional cloning of the gene for X-linked retinitis pigmentosa 2.* Nat Genet, 1998. **19**(4): p. 327-32.

104. Miki, Y., et al., *Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer.* Cancer Res, 1992. **52**(3): p. 643-5.

105. Kines, K.J., et al., *Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci.* Nucleic Acids Res, 2014. **42**(16): p. 10488-502.

106. Wallace, N.A., V.P. Belancio, and P.L. Deininger, *L1 mobile element expression causes multiple types of toxicity.* Gene, 2008. **419**(1-2): p. 75-81.

107. Nikitin, A.G. and R.J. Shmookler Reis, *Role of transposable elements in age-related genomic instability.* Genet Res, 1997. **69**(3): p. 183-95.

108. Bellizzi, D., et al., *Global DNA methylation in old subjects is correlated with frailty.* Age (Dordr), 2012. **34**(1): p. 169-79.

109. Hata, K. and Y. Sakaki, *Identification of critical CpG sites for repression of L1 transcription by DNA methylation.* Gene, 1997. **189**(2): p. 227-34.

110. Chiappinelli, K.B., et al., *Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses.* Cell, 2015. **162**(5): p. 974-86.

111. Ardeljan, D., et al., *Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication.* Nat Struct Mol Biol, 2020. **27**(2): p. 168-178.

112. Tunbak, H., et al., *The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s.* Nat Commun, 2020. **11**(1): p. 5387.

113. Lee, H.J., T.A. Hore, and W. Reik, *Reprogramming the methylome: erasing memory and creating diversity.* Cell Stem Cell, 2014. **14**(6): p. 710-9.

114. Garcia-Perez, J.L., et al., *LINE-1 retrotransposition in human embryonic stem cells.* Hum Mol Genet, 2007. **16**(13): p. 1569-77.

115. Muotri, A.R., et al., *Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.* Nature, 2005. **435**(7044): p. 903-10.

116. Coufal, N.G., et al., *L1 retrotransposition in human neural progenitor cells.* Nature, 2009. **460**(7259): p. 1127-31.

117. Upton, K.R., et al., *Ubiquitous L1 mosaicism in hippocampal neurons.* Cell, 2015. **161**(2): p. 228-39.

118. Marchetto, M.C., et al., *A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells.* Cell, 2010. **143**(4): p. 527-39.

119. Coufal, N.G., et al., *Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells.* Proc Natl Acad Sci U S A, 2011. **108**(51): p. 20382-7.

120. Li, W., et al., *Transposable elements in TDP-43-mediated neurodegenerative disorders.* PLoS One, 2012. **7**(9): p. e44099.

121. Bundo, M., et al., *Increased l1 retrotransposition in the neuronal genome in schizophrenia.* Neuron, 2014. **81**(2): p. 306-13.

122. Lee, W., et al., *MeCP2 regulates gene expression through recognition of H3K27me3.* Nat Commun, 2020. **11**(1): p. 3140.

123. Trivedi, M., et al., *Morphine induces redox-based changes in global DNA methylation and retrotransposon transcription by inhibition of excitatory amino acid transporter type 3-mediated cysteine uptake.* Mol Pharmacol, 2014. **85**(5): p. 747-57.

124. Rusiecki, J.A., et al., *DNA methylation in repetitive elements and post-traumatic stress disorder: a case-control study of US military service members.* Epigenomics, 2012. **4**(1): p. 29-40.

125. Rodic, N., et al., *Long interspersed element-1 protein expression is a hallmark of many human cancers.* Am J Pathol, 2014. **184**(5): p. 1280-6.

126. He, Z.M., et al., *Transition of LINE-1 DNA methylation status and altered expression in first and third trimester placentas.* PLoS One, 2014. **9**(5): p. e96994.

127. Zadora, J., et al., *Disturbed Placental Imprinting in Preeclampsia Leads to Altered Expression of DLX5, a Human-Specific Early Trophoblast Marker.* Circulation, 2017. **136**(19): p. 1824-1839.

128. Ambros, V., *MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing.* Cell, 2003. **113**(6): p. 673-6.

129. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function.* Cell, 2004. **116**(2): p. 281-97.

130. Ansel, K.M., *RNA regulation of the immune system.* Immunol Rev, 2013. **253**(1): p. 5-11.

131. Djuranovic, S., A. Nahvi, and R. Green, *A parsimonious model for gene regulation by miRNAs.* Science, 2011. **331**(6017): p. 550-3.

132. Subramanyam, D. and R. Blelloch, *From microRNAs to targets: pathway discovery in cell fate transitions.* Curr Opin Genet Dev, 2011. **21**(4): p. 498-503.

133. Borchert, G.M., et al., *Comprehensive analysis of microRNA genomic loci identifies pervasive repetitive-element origins.* Mob Genet Elements, 2011. **1**(1): p. 8-17.

134. Piriyapongsa, J. and I.K. Jordan, *A family of human microRNA genes from miniature inverted-repeat transposable elements.* PLoS One, 2007. **2**(2): p. e203.

135. Piriyapongsa, J., L. Marino-Ramirez, and I.K. Jordan, *Origin and evolution of human microRNAs from transposable elements.* Genetics, 2007. **176**(2): p. 1323-37.

136. Smalheiser, N.R. and V.I. Torvik, *Mammalian microRNAs derived from genomic repeats.* Trends Genet, 2005. **21**(6): p. 322-6.

137. Wei, G., et al., *MDTE DB: a database for microRNAs derived from Transposable element.* IEEE/ACM Trans Comput Biol Bioinform, 2016. **13**(6): p. 1155-1160.

138. Qin, S., et al., *The Role of Transposable Elements in the Origin and Evolution of MicroRNAs in Human.* PLoS One, 2015. **10**(6): p. e0131365.

139. Roberts, J.T., S.E. Cardin, and G.M. Borchert, *Burgeoning evidence indicates that microRNAs were initially formed from transposable element sequences.* Mob Genet Elements, 2014. **4**: p. e29255.

140. Iorio, M.V. and C.M. Croce, *Causes and consequences of microRNA dysregulation.* Cancer J, 2012. **18**(3): p. 215-22.

141. Li, Y. and K.V. Kowdley, *MicroRNAs in common human diseases.* Genomics Proteomics Bioinformatics, 2012. **10**(5): p. 246-53.

142. Gu, Y., et al., *Differential miRNA expression profiles between the first and third trimester human placentas.* Am J Physiol Endocrinol Metab, 2013. **304**(8): p. E836-43.

143. Morales-Prieto, D.M., et al., *Pregnancy-associated miRNA-clusters.* J Reprod Immunol, 2013. **97**(1): p. 51-61.

144. Chen, D.B. and W. Wang, *Human placental microRNAs and preeclampsia.* Biol Reprod, 2013. **88**(5): p. 130.

145. Cross, C.E., et al., *Oxidative Stress Alters miRNA and Gene Expression Profiles in Villous First Trimester Trophoblasts.* Biomed Res Int, 2015. **2015**: p. 257090.

146. Choudhury, M. and J.E. Friedman, *Epigenetics and microRNAs in preeclampsia.* Clin Exp Hypertens, 2012. **34**(5): p. 334-41.

147. Zhu, X.M., et al., *Differential expression profile of microRNAs in human placentas from preeclamptic pregnancies vs normal pregnancies.* Am J Obstet Gynecol, 2009. **200**(6): p. 661 e1-7.

148. Stephanie, O., O. Michael, and S. Karolina, *Normal Pregnancy: A Clinical Review.* Academic Journal of Pediatrics and Neonatology, 2016. **1**(1): p. 555554.

149. Maruyama, T. and Y. Yoshimura, *Molecular and cellular mechanisms for differentiation and regeneration of the uterine endometrium.* Endocr J, 2008. **55**(5): p. 795-810.

150. Okada, H., T. Tsuzuki, and H. Murata, *Decidualization of the human endometrium.* Reprod Med Biol, 2018. **17**(3): p. 220-227.

151. Norwitz, E.R., D.J. Schust, and S.J. Fisher, *Implantation and the survival of early pregnancy.* N Engl J Med, 2001. **345**(19): p. 1400-8.

152. Red-Horse, K., et al., *Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface.* J Clin Invest, 2004. **114**(6): p. 744-54.

153. Johnson, M.H. and C.A. Ziomek, *The foundation of two distinct cell lineages within the mouse morula.* Cell, 1981. **24**(1): p. 71-80.

154. Carson, D.D., et al., *Embryo implantation.* Dev Biol, 2000. **223**(2): p. 217-37.

155. Soma-Pillay, P., et al., *Physiological changes in pregnancy.* Cardiovasc J Afr, 2016. **27**(2): p. 89-94.

156. Barker, D.J. and K.L. Thornburg, *Placental programming of chronic diseases, cancer and lifespan: a review.* Placenta, 2013. **34**(10): p. 841-845.

157. Wang, Y. *Vascular biology of the placenta*. in *Colloquium Series on Integrated Systems Physiology: From Molecule to Function*. 2010. Morgan & Claypool Life Sciences.

158. Carter, A.M., *Comparative studies of placentation and immunology in non-human primates suggest a scenario for the evolution of deep trophoblast invasion and an explanation for human pregnancy disorders.* Reproduction, 2011. **141**(4): p. 391-6.

159. Pijnenborg, R., L. Vercruysse, and A.M. Carter, *Deep trophoblast invasion and spiral artery remodelling in the placental bed of the lowland gorilla.* Placenta, 2011. **32**(8): p. 586-91.

160. Benirschke, K., P. Kaufmann, and R. Baergen, *Pathology of the human placenta*. 2006, New York: Springer.

161. Moore, K., T. Persaud, and M. Torchia, *The placenta and fetal membranes. The Developing Human: Clinically Oriented Embryology*. 2008, Philadelphia: Saunders Elsevier Inc.

162. Napso, T., et al., *The Role of Placental Hormones in Mediating Maternal Adaptations to Support Pregnancy and Lactation.* Front Physiol, 2018. **9**: p. 1091.

163. Goplerud, J.M. and M. Delivoria-Papadopoulos, *Physiology of the placenta--gas exchange.* Ann Clin Lab Sci, 1985. **15**(4): p. 270-8.

164. Knipp, G.T., K.L. Audus, and M.J. Soares, *Nutrient transport across the placenta.* Adv Drug Deliv Rev, 1999. **38**(1): p. 41-58.

165. Macias, R.I., J.J. Marin, and M.A. Serrano, *Excretion of biliary compounds during intrauterine life.* World J Gastroenterol, 2009. **15**(7): p. 817-28.

166. PrabhuDas, M., et al., *Immune mechanisms at the maternal-fetal interface: perspectives and challenges.* Nat Immunol, 2015. **16**(4): p. 328-34.

167. Burton, G.J. and A.L. Fowden, *The placenta: a multifaceted, transient organ.* Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1663): p. 20140066.

168. Golightly, E., H.N. Jabbour, and J.E. Norman, *Endocrine immune interactions in human parturition.* Mol Cell Endocrinol, 2011. **335**(1): p. 52-9.

169. Turco, M.Y. and A. Moffett, *Development of the human placenta.* Development, 2019. **146**(22).

170. Knofler, M., et al., *Human placenta and trophoblast development: key molecular mechanisms and model systems.* Cell Mol Life Sci, 2019. **76**(18): p. 3479-3496.

171. Kingdom, J., et al., *Development of the placental villous tree and its consequences for fetal growth.* Eur J Obstet Gynecol Reprod Biol, 2000. **92**(1): p. 35-43.

172. Caniggia, I., et al., *Inhibition of TGF-beta 3 restores the invasive capability of extravillous trophoblasts in preeclamptic pregnancies.* J Clin Invest, 1999. **103**(12): p. 1641-50.

173. Carlson, B., *Human Embryology and Developmental Biology*

5th Edition. 2012: Elsevier.

174. Saleem, S., et al., *A prospective study of maternal, fetal and neonatal deaths in low- and middle-income countries.* Bull World Health Organ, 2014. **92**(8): p. 605-12.

175. Steegers, E.A., et al., *Pre-eclampsia.* Lancet, 2010. **376**(9741): p. 631-44.

176. Robillard, P.Y., et al., *Preeclampsia and human reproduction. An essay of a long term reflection.* J Reprod Immunol, 2003. **59**(2): p. 93-100.

177. Tranquilli, A.L., et al., *The classification, diagnosis and management of the hypertensive disorders of pregnancy: A revised statement from the ISSHP.* Pregnancy Hypertens, 2014. **4**(2): p. 97-104.

178. von Dadelszen, P., L.A. Magee, and J.M. Roberts, *Subclassification of preeclampsia.* Hypertens Pregnancy, 2003. **22**(2): p. 143-8.

179. Duley, L., *The global impact of pre-eclampsia and eclampsia.* Semin Perinatol, 2009. **33**(3): p. 130-7.

180. Organization, W.H., *WHO recommendations for prevention and treatment of pre-eclampsia and eclampsia.* 2011.

181. English, F.A., L.C. Kenny, and F.P. McCarthy, *Risk factors and effective management of preeclampsia.* Integr Blood Press Control, 2015. **8**: p. 7-12.

182. Roberts, J.M. and D.W. Cooper, *Pathogenesis and genetics of pre-eclampsia.* Lancet, 2001. **357**(9249): p. 53-6.

183. Robillard, P.Y., et al., *Association of pregnancy-induced hypertension with duration of sexual cohabitation before conception.* Lancet, 1994. **344**(8928): p. 973-5.

184. Saito, S., et al., *Inadequate tolerance induction may induce pre-eclampsia.* J Reprod Immunol, 2007. **76**(1-2): p. 30-9.

185. Dekker, G., *The partner's role in the etiology of preeclampsia.* J Reprod Immunol, 2002. **57**(1-2): p. 203-15.

186. Redman, C.W., *Current topic: pre-eclampsia and the placenta.* Placenta, 1991. **12**(4): p. 301-8.

187. Redman, C.W. and I.L. Sargent, *Immunology of pre-eclampsia.* Am J Reprod Immunol, 2010. **63**(6): p. 534-43.

188. Robertson, S.A., et al., *Activating T regulatory cells for tolerance in early pregnancy - the contribution of seminal fluid.* J Reprod Immunol, 2009. **83**(1-2): p. 109-16.

189. Robertson, S.A., et al., *Seminal fluid drives expansion of the CD4+CD25+ T regulatory cell pool and induces tolerance to paternal alloantigens in mice.* Biol Reprod, 2009. **80**(5): p. 1036-45.

190. Vinatier, D. and J.C. Monnier, *Pre-eclampsia: physiology and immunological aspects.* Eur J Obstet Gynecol Reprod Biol, 1995. **61**(2): p. 85-97.

191. Saito, S. and M. Sakai, *Th1/Th2 balance in preeclampsia.* J Reprod Immunol, 2003. **59**(2): p. 161-73.

192. Moffett, A. and S.E. Hiby, *How Does the maternal immune system contribute to the development of pre-eclampsia?* Placenta, 2007. **28 Suppl A**: p. S51-6.

193. Rajagopalan, S. and E.O. Long, *Understanding how combinations of HLA and KIR genes influence disease.* J Exp Med, 2005. **201**(7): p. 1025-9.

194. Phipps, E.A., et al., *Pre-eclampsia: pathogenesis, novel diagnostics and therapies.* Nat Rev Nephrol, 2019. **15**(5): p. 275-289.

195. Ahmed, A., et al., *Induction of placental heme oxygenase-1 is protective against TNFalpha-induced cytotoxicity and promotes vessel relaxation.* Mol Med, 2000. **6**(5): p. 391-409.

196. Hecht, J.L., et al., *The pathology of eclampsia: An autopsy series.* Hypertens Pregnancy, 2017. **36**(3): p. 259-268.

197. Powe, C.E., R.J. Levine, and S.A. Karumanchi, *Preeclampsia, a disease of the maternal endothelium: the role of antiangiogenic factors and implications for later cardiovascular disease.* Circulation, 2011. **123**(24): p. 2856-69.

198. Levine, R.J., et al., *Soluble endoglin and other circulating antiangiogenic factors in preeclampsia.* N Engl J Med, 2006. **355**(10): p. 992-1005.

199. Zeisler, H., et al., *Predictive Value of the sFlt-1:PlGF Ratio in Women with Suspected Preeclampsia.* N Engl J Med, 2016. **374**(1): p. 13-22.

200. Burke, S.D., et al., *Soluble fms-like tyrosine kinase 1 promotes angiotensin II sensitivity in preeclampsia.* J Clin Invest, 2016. **126**(7): p. 2561-74.

201. Dechend, R., et al., *AT(1) receptor agonistic antibodies from preeclamptic patients cause vascular cells to express tissue factor.* Circulation, 2000. **101**(20): p. 2382-7.

202. Dechend, R., et al., *AT1 receptor agonistic antibodies from preeclamptic patients stimulate NADPH oxidase.* Circulation, 2003. **107**(12): p. 1632-9.

203. Zhou, C.C., et al., *Angiotensin receptor agonistic autoantibodies induce pre-eclampsia in pregnant mice.* Nat Med, 2008. **14**(8): p. 855-62.

204. Hubel, C.A., et al., *Agonistic angiotensin II type 1 receptor autoantibodies in postpartum women with a history of preeclampsia.* Hypertension, 2007. **49**(3): p. 612-7.

205. Wang, A., S. Rana, and S.A. Karumanchi, *Preeclampsia: the role of angiogenic factors in its pathogenesis.* Physiology (Bethesda), 2009. **24**: p. 147-58.

206. Merlo, G.R., et al., *Multiple functions of Dlx genes.* Int J Dev Biol, 2000. **44**(6): p. 619-26.

207. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.

208. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* Elife, 2015. **4**.

209. Goodier, J.L., et al., *A potential role for the nucleolus in L1 retrotransposition.* Hum Mol Genet, 2004. **13**(10): p. 1041-8.

210. Price, E.M., et al., *Different measures of "genome-wide" DNA methylation exhibit unique properties in placental and somatic tissues.* Epigenetics, 2012. **7**(6): p. 652-63.

211. Rodic, N. and K.H. Burns, *Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms?* PLoS Genet, 2013. **9**(3): p. e1003402.

212. Portella, G., F. Battistini, and M. Orozco, *Understanding the connection between epigenetic DNA methylation and nucleosome positioning from computer simulations.* PLoS Comput Biol, 2013. **9**(11): p. e1003354.

213. Dupressoir, A., et al., *Syncytin-A knockout mice demonstrate the critical role in placentation of a fusogenic, endogenous retrovirus-derived, envelope gene.* Proc Natl Acad Sci U S A, 2009. **106**(29): p. 12127-32.

214. Sokolowski, M., et al., *Truncated ORF1 proteins can suppress LINE-1 retrotransposition in trans.* Nucleic Acids Res, 2017. **45**(9): p. 5294-5308.

215. Freeman, B.T., et al., *Identification of charged amino acids required for nuclear localization of human L1 ORF1 protein.* Mob DNA, 2019. **10**: p. 20.

216. Goodier, J.L., et al., *LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex.* Mol Cell Biol, 2007. **27**(18): p. 6469-83.

217. Ostertag, E.M., et al., *Determination of L1 retrotransposition kinetics in cultured cells.* Nucleic Acids Res, 2000. **28**(6): p. 1418-23.

218. Ardeljan, D., et al., *LINE-1 ORF2p expression is nearly imperceptible in human cancers.* Mob DNA, 2020. **11**: p. 1.

219. Sanchez-Luque, F.J., S.R. Richardson, and G.J. Faulkner, *Retrotransposon Capture Sequencing (RC-Seq): A Targeted, High-Throughput Approach to Resolve Somatic L1 Retrotransposition in Humans.* Methods Mol Biol, 2016. **1400**: p. 47-77.

220. Wildman, D.E., et al., *Evolution of the mammalian placenta revealed by phylogenetic analysis.* Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3203-8.

221. Elliot, M.G. and B.J. Crespi, *Phylogenetic evidence for early hemochorial placentation in eutheria.* Placenta, 2009. **30**(11): p. 949-67.

222. Mess, A. and A.M. Carter, *Evolutionary transformations of fetal membrane characters in Eutheria with special reference to Afrotheria.* J Exp Zool B Mol Dev Evol, 2006. **306**(2): p. 140-63.

223. Hou, Z., et al., *Adaptive history of single copy genes highly expressed in the term human placenta.* Genomics, 2009. **93**(1): p. 33-41.

224. Older Aguilar, A.M., et al., *Coevolution of killer cell Ig-like receptors with HLA-C to become the major variable regulators of human NK cells.* J Immunol, 2010. **185**(7): p. 4238-51.

225. Hu, S., et al., *SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation.* PLoS Genet, 2015. **11**(7): p. e1005367.

226. Kopera, H.C., et al., *LINE-1 Cultured Cell Retrotransposition Assay.* Methods Mol Biol, 2016. **1400**: p. 139-56.

227. Borzychowski, A.M., I.L. Sargent, and C.W. Redman, *Inflammation and pre-eclampsia.* Semin Fetal Neonatal Med, 2006. **11**(5): p. 309-16.

228. Redman, C.W. and I.L. Sargent, *Preeclampsia and the systemic inflammatory response.* Semin Nephrol, 2004. **24**(6): p. 565-70.

229. Rein, D.T., et al., *Preeclamptic women are deficient of interleukin-10 as assessed by cytokine release of trophoblast cells in vitro.* Cytokine, 2003. **23**(4-5): p. 119-25.

230. Sakai, M., et al., *The ratio of interleukin (IL)-18 to IL-12 secreted by peripheral blood mononuclear cells is increased in normal pregnant subjects and decreased in pre-eclamptic patients.* J Reprod Immunol, 2004. **61**(2): p. 133-43.

231. Benton, S.J., et al., *The clinical heterogeneity of preeclampsia is related to both placental gene expression and placental histopathology.* Am J Obstet Gynecol, 2018. **219**(6): p. 604 e1-604 e25.

232. De Cecco, M., et al., *L1 drives IFN in senescent cells and promotes age-associated inflammation.* Nature, 2019. **566**(7742): p. 73-78.

233. Simon, M., et al., *LINE1 Derepression in Aged Wild-Type and SIRT6-Deficient Mice Drives Inflammation.* Cell Metab, 2019. **29**(4): p. 871-885 e5.

234. Yu, Q., et al., *DNA-damage-induced type I interferon promotes senescence and inhibits stem cell function.* Cell Rep, 2015. **11**(5): p. 785-797.

235. Gasior, S.L., et al., *The human LINE-1 retrotransposon creates DNA double-strand breaks.* J Mol Biol, 2006. **357**(5): p. 1383-93.

236. Zhang, T., et al., *Dissecting human trophoblast cell transcriptional heterogeneity in preeclampsia using single-cell RNA sequencing.* Mol Genet Genomic Med, 2021. **9**(8): p. e1730.

237. De Donato, M., et al., *Conservation of Repeats at the Mammalian KCNQ1OT1-CDKN1C Region Suggests a Role in Genomic Imprinting.* Evol Bioinform Online, 2017. **13**: p. 1176934317715238.

238. Cobo, T., et al., *Role of Periostin in Adhesion and Migration of Bone Remodeling Cells.* PLoS One, 2016. **11**(1): p. e0147837.

239. Kabir-Salmani, M., et al., *Alphavbeta3 integrin signaling pathway is involved in insulin-like growth factor I-stimulated human extravillous trophoblast cell migration.* Endocrinology, 2003. **144**(4): p. 1620-30.

240. Gillan, L., et al., *Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility.* Cancer Res, 2002. **62**(18): p. 5358-64.

241. Ahn, H.W., et al., *Progesterone and interferon tau-regulated genes in the ovine uterine endometrium: identification of periostin as a potential mediator of conceptus elongation.* Reproduction, 2009. **138**(5): p. 813-25.

242. Freis, A., et al., *Serum periostin levels in early in pregnancy are significantly altered in women with miscarriage.* Reprod Biol Endocrinol, 2017. **15**(1): p. 87.

243. Li, J., B. LaMarca, and J.F. Reckelhoff, *A model of preeclampsia in rats: the reduced uterine perfusion pressure (RUPP) model.* Am J Physiol Heart Circ Physiol, 2012. **303**(1): p. H1-8.

244. Ishida, J., et al., *Pregnancy-associated homeostasis and dysregulation: lessons from genetically modified animal models.* J Biochem, 2011. **150**(1): p. 5-14.

245. Gatford, K.L., et al., *Animal Models of Preeclampsia: Causes, Consequences, and Interventions.* Hypertension, 2020. **75**(6): p. 1363-1381.

246. Ozdemir, O., *Various effects of different probiotic strains in allergic disorders: an update from laboratory and clinical data.* Clin Exp Immunol, 2010. **160**(3): p. 295-304.

7. ABBREVIATIONS:

A = adenine
AngII = angiotensin II
APOBEC = apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
ATAC-seq = assay for Transposase-Accessible Chromatin using sequencing
casTLE = Cas9 high-Throughput maximum Likelihood Estimator
cDNA = copy DNA
CNV = copy number variation
CPM = count per million
ddH_2O = double distilled H_2O
DLX5 = distal-less homeobox 5
DNA = deoxyribonucleic acid
DNMT = DNA methyltransferase
EGF = epidermal growth factor
EN = endonuclease
EOPE = early onset preeclampsia
GA = gestational age
H3K9me3 = histone H3 lysine 9 trimethylation
HO = heam oxygenase
hsCRP = high-sensitivity C-reactive protein
HIF1 = hypoxia-inducible factor 1
HUSH complex = human silencing hub complex
ICM = inner cell mass
IF = immunofluorescence
IGF-1 = insulin growth factor-1
IHC = immunohistochemistry
INDELS = insertions / deletions
iPSCs = induced pluripotent stem cells
IUGR = intra-uterine growth restriction
kDA = kilodalton
KIRs = killer immunoglobulin-like receptors
KRAB = Kruppel-associated box
KRAB-ZFPs = KRAB-zinc finger proteins
L1HS = human specific LINE-1
LH = luteinizing hormone
LINE = long interspersed nuclear element
LNA = locked nucleic acid
LOPE = late onset preeclampsia
LTR = long terminal repeats
MeCP2 = methyl CPG binding protein 2
miRNA = microRNA
NK cells = natural killer cells
NO = nitric oxide
OE = over expressing

ORF = open reading frame
PCA = principal component analysis
PDPK = proline-directed protein kinase
PE = preeclampsia
PFA = paraformaldehyde
piRNAs = PIWI RNAs
PlGF = placental growth factor
PolyA = polyadenylation
POSTN = periostin
RC-seq = retrotransposon-capture sequencing
RISC = RNA induced silencing complex
RNA = ribonucleic acid
RNAi = RNA interference
RNP = ribonucleoprotein particle
ROS = reactive oxygen species
RT = reverse transcriptase
sEng = soluble endoglin
sFLT = soluble Fms-like thyrosine kinase
SINE = short interspersed nuclear element
SVA = SINE-R/VNTR/Alu-like retrotransposons
T = thymine
TE = transposable element
TGFβ = Transforming growth factor beta
Th1/2 = T helper type 1/2
TNFα = tumor necrosis factor
TPM = transcript per million
TPRT = target primed reverse transcription
TrE = trophectoderm
Tregs – regulatory T cells
TSD = target site duplication
TSH = thyroid stimulating hormone
VEGF = vascular endothelial growth factor
WT = wildtype

8. ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisors Dr. Zsuzsanna Izsvák and Prof. Dr. Ralf Dechend, for giving me the opportunity to conduct my PhD research in their labs on this fascinating fusion project about reproductive and molecular biology. I would also like to thank Prof. Dr. Petra Knaus for her support as an Academic Supervisor in Freie Universität Berlin and for reviewing my thesis.

In addition, I would like to thank my former mentors Prof. Dr. Imogen Coe and the Coe Lab from Toronto, Canada, and Prof. Dr. Nicholas Gompel from LMU, Munich, Germany, who have inspired me to pursue a scientific career and who have molded me into the scientist that I am today. You've taught me to be brave, poised, precise, professional, and most of all, how to be a good scientist. I still can hear your words of guidance today, "do not be apologetic" and "never do an experiment for the sake of using a technique, do it to answer a question." Another big thank you to my most recent supervisors, Dr. Petra Zalud and Prof. Dr. Torsten Zuberbier, for being so understanding and supportive during the final phases of my PhD and for propelling me forward in my career.

Thank you to Dr. Florian Herse from AG Dechend-Müller for the immense support he has given me throughout my PhD. I sincerely thank you for the scientific help but also the counseling and friendship. Also, I would like to greatly thank Jule Anders for the technical support, I have learned a lot from you, and thank you for all the wonderful talks.

I would like to thank all my former and current AG Izsvak and MDC colleagues. Thank you to Kathy, Sasha, Felix, Julia, Ilija, Angelika (and Juno), Agniezska, Yuliang, Himanshu, Christine, Rabia, Sandra, Bea, Bertrand, Chigoze, Amit, Guo, Zhimin, Karam, Tamas, Cai, and Manu. Your friendship has made this experience so much more special. I will cherish dearly all the laughs, bike rides, car rides, and lake trips that we've had together. A special thanks to all whom I could always count on to change my media when I could not make it to the MDC.

Thank you to all my old friends who have been rooting for me since day one. Thank you to my life-long best friend Nevena Prekic, my YorkU partners in crime: Kate of K^2 with whom I share the experience of moving to Germany, and the amazing Dr. Mons, my first friends in Canada Nina and Nadja, my wonderful friend Anja, and the fantastic friends that I've met in Berlin: Jelena – my personal mentor, Tanja, Adela, and Aleks.

And a very special thanks to Felix, Sasha, Kathy, and Jelena, who took the time to proofread my thesis!

Finally, I would like to thank my wonderful inner and outer family who has given me so much support that a person can do anything. Thank you Nana Novka- your care and support made my student life much easier, Deka Branko for the life lessons, Teta Nena and family, Ujka and Danica and family, and Tetka Zorica and family. A special thanks to my talented cousin Jelena who helped with illustrations of my figures. Thank you to my grandma Rada who wished me to keep going to school for many more years to come, the last time that I've ever spoken with her. Thank you to my late grandpas Sava and Siniša, the thought of them warms my heart. I endlessly thank my mom Marina and dad Dragan, the most supportive and loving parents one could ask for. They have always supported my decisions, emotions, and ambitions. They have let me pave my path and are always behind me every step of the way. They speed drove to bring me a calculator for my exam, made snacks while I studied, trusted me to come to Munich for my master's studies, and proudly supported me during my PhD. For everything that has happened until now, they've shared with me every pain and joy. Mom and dad, I couldn't have done it without you. A special thank you to my sister the better scientist, Jasmina, who is why I chose the scientific career. Dado, I've always admired the beautiful, interesting, smart, and fun person you are, and I wanted to be just like you. Thank you for being my role model and for your endless support. Thank you to my brother-in-law Dušan for always asking how I am doing and wishing me luck. Thank you to my nephew Filip for melting my heart and brightening up my day with every picture. Thank you to my Gare, who kept my feet warm while studying since my bachelor studies, whose picture was the background of my desktop for my entire PhD that put a smile on my face, and who waits with love and happiness for me to come home and visit. Thank you to Ozi, who cuddled and purred to comfort me. I love and appreciate you all.

9. Curriculum Vitae of Katarina Stevanovic

Education:

- Doctor of Philosophy, Biochemistry 2018 - Present
  Free University, Berlin, Germany
- Master of Science, Biology 2013 - 2015
  Ludwig - Maximilians University, Munich, Germany
- Bachelor of Science, Specialized Honours in Biochemistry 2008 - 2013
  York University, Toronto, Ontario, Canada

Experience:

- **Scientific Writer** **March 2021 – Present**
  Workgroup of Prof. Dr. Torsten Zuberbier, Charité Medical University, Berlin, Germany
- **Scientific Project Manager** **June 2019 – January 2021**
  TP21, Berlin, Germany
- **Biology Lecturer** **February 2020 – June 2020**
  Touro College, Berlin, Germany
- **Research Scientist – PhD Thesis** **May 2016 – September 2021**
  Collaboration of Max-Delbrück Centre for Molecular Medicine (MDC) and Charité Medical University, Berlin, Germany
  Workgroups of Dr. Zsuzsanna Izsvak, (MDC); Prof. Dr. Ralf Dechend (Charité)
  Project: LINE-1 elements in human placenta and their role in preeclampsia.
- **Research Visitor** **February – April 2015**
  Collaboration of St. Michael's Hospital and Ryerson University, Toronto, ON, Canada
  Work Group of Prof. Dr. Imogen Coe
  Project: ENT2 protein knock-down in HEK293T cells.
- **Molecular Biology Techniques Teaching Assistant** **October 2014, October 2015**
  Ludwig-Maximilians University (LMU), Munich, Germany
- **Working Student** **2014 - 2015**
  Ludwig-Maximillians University (LMU), Munich, Germany
  Workgroup of Prof. Dr. Nicolas Gompel
  Project: GFP tagging of various genes in *D. melanogaster* by CRISPR/Cas genome editing.

Courses:

- Statistics in R
- GraphPad Prism
- Adobe Illustrator
- Career Development
- Professional Communication
- Project Management
- Scientific Writing & Publishing
- Web Design (HTML, CSS, Java Script)

Publications:

- **Stevanovic, K**. et al. LINE-1 elements in human placenta and PE pathogenesis. In prep.
- Zuberbier T, …**Stevanovic, K**, et al. (2021) Proposal of 0.5mg of protein /100g of processed food as threshold for voluntary declaration of food allergen traces in processed food – a first step in an initiative to better inform patients and avoid fatal allergic reactions - A GA²LEN position paper. *Allergy,* Submitted.
- Hofmann MA, Fluhr JW, Ruwwe-Glösenkamp C, **Stevanovic K**, Bergmann K, Zuberbier T. (2021) Role of IL-17 in atopy – a systematic review. *Clinical and Translational Allergy*, 11(6).
- Grañe-Boladeras N, Williams D, Tarmakova Z, **Stevanovic K**, Villani L, Mehrabi P, Siu KWM, Pastor-Anglada M, and Coe IR. (2018) Oligomerization of Equilibrative Nucleoside Transporters: a novel regulatory and functional mechanism involving PKC and PP1. *FASEB J*.
- Ramadan A, Naydenova Z, **Stevanovic K**, Rose J B., and Coe IR. (2014). The adenosine transporter, ENT1, in cardiomyocytes is sensitive to inhibition by ethanol in a kinase-dependent manner: implications for ethanol-dependent cardioprotection and nucleoside analog drug cytotoxicity. *Purinergic Signalling*, *10*(2), 305–312.

Awards:

- Best Poster Presentation – Magdeburg Autumn School on Reproductive Biology 2017
- Independent Researcher - Lehre@LMU Award 2014

Conferences:

- 2nd International Conference on the Long and Short Non-Coding RNAs 9-14 June 2017, Crete, Greece
- Transcard Retreat, Valencia, Spain
- Trophoblast Conference, Cambridge, UK
- Interdisciplinary Autumn School for Reproductive Sciences and Related Research Fields 18-21 October 2017, Magdeburg, Germany
- CMD Retreat – A systems approach to cardiovascular and metabolic diseases, September 18-21, Bad Sarrow, Germany