

Roles of Endogenous Retroviruses in Early Life Events

A retrovirus that infected our ancestors 100 million years ago became a human gene that is expressed in embryos and cancers, and can be detected in the blood of pregnant women. Accumulating evidence suggests potential roles for endogenous retroviruses in early life events, which may affect adult health.

The discovery of reverse transcriptase in 1970 shattered our DNA-to-RNA central dogma for one-way flow of biological information. This reverse genetic flow has facilitated the colonization of vertebrate genomes by retroviruses, a process that we know has been active for at least 450 million years [1]. Known as endogenous retroviruses (ERVs), these abundant genomic parasites comprise 8% of the human genome.

What do these efficient genomic colonizers do? Are they merely fossils that, like mosquitos in amber, were stuck and preserved in large host genomes while their functions decayed? Researchers have been struggling to understand their roles for as long as we have known them, postulating junk, bystander, and pathogen hypotheses. Research in this area started when retroviruses were accidentally discovered in the early 1900s as unseen transmissible agents of cancer. The strong links of retroviruses with animal cancers created one great promise: if researchers identified infectious agents causing human cancer, they could design prevention strategies (e.g., vaccines) or drug treatments for cancer. Although the vast majority of human cancers have not been linked with retroviruses (with the notable exceptions of human T lymphotropic virus, HTLV, and HIV) [2], by studying retroviruses we understand cancer biology much better than before. Not more than 15–20 years ago, research into human retroviral cancers was counting a high dead-end project toll; it was the time when unfruitful hunting for potential human retroviral cancers lead to the term ‘rumor viruses’ over the hope for discovering actual ‘tumor viruses’.

In the meantime, major breakthroughs on potential roles of co-opted retroviruses in physiology were achieved. Around the time that the first draft of the human genome was published, two groups identified two tamed retroviral envelope genes with functionality in the human placenta [3]. In the years that followed such placental retroviral-envelope genes were recognized in diverse mammals, including marsupials. These tamed envelopes offer two functions for their host. First, they promote fusion of cells in placentas, which is important for the anatomy of the syncytiotrophoblast (hence the name syncytins). Second, they protect the embryo from the mother’s immunity (fetomaternal immunotolerance).

Heidmann’s group has recently identified a new tamed retroviral envelope that is produced from the fetus and then shed in the mother’s blood during pregnancy [4]. It is named HEMO (human endogenous

MER34 ORF) and is the first retroviral envelope described to be shed extracellularly, even though similar shedding has been described in envelopes of other viruses (e.g., Ebola) [4]. The functionality of this tamed envelope remains unclear. It has low expression in a variety of normal tissues, but is highly expressed in the placenta, a variety of stem cells (including neural and induced pluripotent), and cancer tissues (most strikingly ovarian cancer). During the development of the embryo, expression of HEMO peaks at later stages compared to the expression of HERV-K HML-2 [4, 5], the human ERV (HERV) with the most recent integrations in the human genome. In support of an important functional role, we see the characteristic evolutionary signature: HEMO has been extraordinarily preserved over the last 100 million years. This strong purifying selection, combined with the expression pattern, suggests an important functional role for HEMO in stem cells and early life events in primates. In contrast, HEMO deteriorated in other nonprimate mammals, suggesting that its functionality was made redundant.

Similarly to HEMO, many ERVs and ERV-derived genes are expressed during the development of embryos [5]. This pattern probably echoes a strategy for expansion in the germlines: expression and copying within stem cells allows proliferation and expansion within cell lineages of the host. Copying in differentiated cells is a proviral dead-end strategy, which can only serve production of virions and transmission between hosts, thus the expansion of the proviral state in the germline serves the survival of the virus in deep time. Indeed, accumulating evidence suggests that ERVs are manipulators of stem-cell biology; most strikingly, HERV-H controls stem-cell identity in humans and closely related primates [5].

HEMO's expression pattern suggests potential roles for manipulation of stem cells and early life events, which could have very important effects on adult diseases. For example, epidemiological studies on the associations of early life events with adult diseases have shown that increased birth weight, which is now known to be a surrogate of stem-cell numbers in adult life [6], is associated with lower cardiovascular risk [2]. Crucially, the number of circulating stem cells is thought to lower cardiovascular risk by taking part in the reconstruction and repair of endothelial injury [7].

In contrast, increased birth weight has been associated with a higher risk of breast (and other) cancers in adult life [2]. In 1990, Trichopoulos pinpointed the significance of early events in adult cancer by asking 'Does breast cancer originate *in utero*?' [8] and a few years later 'Is cancer causation simpler than we thought, but more intractable?' [9]. He was mostly interested in breast cancer as a model to study early life events and their connections with adulthood cancer: a proportion of breast stem cells remain undifferentiated until the first full-term pregnancy. This model explains why earlier pregnancies protect from breast cancer: breast stem cells are forced to differentiate earlier, thus they are less likely to cause cancer.

The correlation between number of stem cells at birth and cancer, when extrapolated among different species, predicts that larger animals should die of cancer much younger, yet they do not. This paradox was pinpointed by Sir Richard Peto [10] and can be resolved through evolutionary pressure that operates over millions of years. Indeed, larger mammals tend to have lower ERV activity, suggesting a deep-time evolutionary pressure to limit the expansion of ERVs [10].

Finally, the manipulation of stem-cell features by ERVs may also contribute to tumor cell dynamics: cancer stem cells have important roles in cancer dynamics. We thus hypothesize that the level of ERV expression in cancers reflects the degree of ERV-initiated cancer stem-cell manipulation as well as the contribution of early life events in the observed pathophysiology. Such roles should be explored to identify potential anticancer treatments.