Marjory Stephenson (1885–1948) was a renowned bacteriologist and biochemist who became the first woman, alongside the chemist Kathleen Lonsdale, to be elected a Fellow of the Royal Society in 1945, whereupon the University of Cambridge belatedly appointed her Reader in Chemical Microbiology. In 1947, she succeeded Alexander Fleming to become the second President of the Society for General Microbiology, but sadly died of cancer at the end of the following year. Since 1953, the Society has commemorated her through the Marjory Stephenson Prize Lecture awarded to recognize ‘exceptional contributions to the discipline of microbiology’. It is a great privilege to deliver the 2015 Lecture, in which I shall discuss selected topics from my research career in the light of what we know today. I remain mindful that several of the salient early discoveries about retroviruses came from phenomena already established a decade earlier for bacteriophages, such as integration into the host genome, transduction of host genes and phenotypic mixing (pseudotypes), though not reverse transcription.

Retroviral pseudotypes

When I was a doctoral student in Michael Abercrombie’s cell biology laboratory in the 1960s, my thesis project was to investigate the loss of contact inhibition of cell proliferation and locomotion in chick embryo fibroblasts by transforming them into tumour cells with Rous sarcoma virus (RSV). Peyton Rous discovered the virus in 1911, but it was the development of a quantitative ‘focus’ assay of transformed cells (Fig. 1a, b) (Temin & Rubin, 1958) that led to a renaissance of RNA tumour virus research (Weiss, 2011). Hanafusa et al. (1963) found that infectious stocks of RSV contained a mixture of two viruses, the transforming virus that was replication-defective, and a replication-competent ‘helper’ avian leukemia virus (ALV). We now know that the replication-defective RSV has an oncogene, src, in place of the env gene of ALV, which provides the missing envelope (Fig. 1c). Rubin called the complemented RSV a ‘pseudotype’ because its neutralization properties and receptor usage were determined by the envelope of the helper virus and not by the genome within the virus particle (Rubin, 1965).

Today’s retroviral pseudotypes are deliberately constructed to carry reporter genes encoding proteins such as luciferase or GFP, but src originally served the same purpose. Exploiting pseudotype viruses has been a recurring method during my career, from discerning endogenous retroviruses in the 1960s (Weiss 1969a, b) to investigating viral entry and neutralizing antibodies. We used vesicular stomatitis virus pseudotypes (Závada, 1976) to detect the expression of avian endogenous retrovirus (ERV) envelopes (Love & Weiss, 1974) and later the HIV-1 envelope to reveal that CD4 is the binding receptor (Dalgleish et al., 1984), and that it is necessary but not sufficient for entry (Maddon et al., 1986), leading to the search for co-receptors. More recently, we used HIV-based lentiviral vectors pseudotyped with 100 envelopes of different HIV-1 subtypes and strains to screen broadly neutralizing single chain antibodies from llamas (McCoy et al., 2012). Pseudotypes of gamma-retroviruses or lentiviruses bearing envelope glycoproteins of highly pathogenic enveloped viruses such as SARS coronavirus (Temperton et al., 2005), H5N1 influenza virus (Temperton et al., 2007), rabies and lyssaviruses (Wright et al., 2008) and Ebola filovirus (Long et al., 2015) have provided sensitive and specific neutralization serology of patients and vaccine recipients without the need for a high containment laboratory.
Retroviral oncogenes

The src oncogene became the prototype oncogene (onc) for many others discovered in acutely transforming RNA tumour viruses in chickens, mice, rats and cats, such as abl, myc and ras (Vogt, 2012). Almost all of these oncogene-bearing retroviruses were replication-defective, although the oncogene was not usually substituted for the env gene; several represent read-through fusion genes from gag into onc (Rosenberg & Jolicoeur, 1997). However, some strains of RSV are replication-competent, although it is likely that the original strain was replication-defective (Weiss & Vogt, 2011), like newly generated avian sarcoma viruses (Geryk et al., 1989). The non-defective RSV strains have the proviral genome structure LTR-gag-pol-env-src-LTR. They enabled the discovery of genetic sequences in the RNA genome of the virion associated with cell transformation (Duesberg & Vogt, 1970) and of genetic recombination between retroviruses in dual infection by RSV and ALV (Vogt, 1971; Weiss et al., 1973).

If retroviral oncogenes are superfluous for viral replication, where did they come from? The provenance of src was solved by DNA annealing experiments (Stehelin et al., 1976) as being a host gene that had been hijacked and mutated by the retrovirus. Other oncogenes had a similar history and appear to have escaped from the host as aberrant RNA transcripts with unusual splicing between the provirus and a splice acceptor in the host (Vogt, 2012).

Leukaemogenic retroviruses lacking oncogenes integrate at near-random sites mainly in ‘open’ regions of host chromosomal DNA. Among the millions of cells infected, some proviruses integrate next to host proto-oncogenes and activate their expression, and the transduction of cellular genes can begin in this way (Vogt, 2012). Moreover, complex recombination events occur between infectiously transmitted retroviruses and endogenous related genomes before the proximate retrovirus activates cellular oncogenes (Rosenberg & Jolicoeur, 1997). For example, in cats, feline leukemia virus subtype A (FeLV-A) is the major transmissible agent, but in most cats that develop leukaemia, recombinination with an endogenous retroviral genome encoding a subtype B env must occur (Roy-Burman, 1995). Multiple interactions occur between exogenous and endogenous LTRs and env regions, giving rise to variants with different pathogenic attributes (Neil et al., 1991; Stewart et al., 2011).

There appears to be little selective advantage to the virus to carry onc, other than that tumour cells tend to be more permissive to retrovirus replication. In fact, there is scant evidence that onc-bearing viruses are naturally transmitted from host to host unless they come to the attention of pathologists and virologists like Rous who deliberately propagate them. The exception is the cyclin oncogene of replication-competent epsilon-retroviruses in fish, but it is probably not of host origin (Rovnak & Quackenbush, 1998).

Fig. 1. The ‘focus’ assay of transformed cells and genome of Rous sarcoma virus. (a) 10-fold serial dilutions of virus on chick embryo fibroblasts showing foci of cell transformation. (b) Scanning electron micrograph of one focus of transformed cells. (c) DNA genomes of replication-defective Rous sarcoma virus (RSV) and replication-competent ‘helper’ virus, avian leukemia virus (ALV), showing the long terminal repeats (LTR) and the open reading frames for the genes essential for replication, gag, pol and env, and the oncogene, src. (a) and (b) reproduced from Weiss (1969c). Bar, 50 μm.
2010). Thus, retroviruses bearing onc genes are rare side effects of infection with little or no bearing on the evolutionary dynamics between virus and host. Nonetheless, the impact of studying such oncogenic retroviruses has been immense because the majority of oncogenes found to be activated in non-viral tumours, including many human cancers, were first identified in animal retroviruses.

In contrast to oncogene-bearing retroviruses, large genome DNA viruses carry many genes originally derived from their hosts. Throughout virus–host evolution, viruses possessing a substantial packaging capacity have incorporated host genes and modified them to serve viral functions such as replication and immune evasion (Knipe & Howley, 2013). However, this was unknown at the time that oncogenes were discovered.

**Endogenous retroviruses (ERVs)**

I have discussed the discovery of endogenous retroviral genomes elsewhere (Weiss, 2006). In this Marjory Stephenson Lecture, I wish to remark that when Jim Payne and I accrued evidence for Mendelian inheritance of retrovirus markers in normal chicken cells – before the discovery of reverse transcriptase in 1970 – only the SGM’s *Journal of General Virology* (Payne & Chubb, 1968; Weiss, 1969a, b) was sufficiently open-minded to publish our findings! Spontaneous production of a replication-competent virus (RAV-0) with similar envelope properties to the ERV that we studied was soon detected in another chicken breed (Vogt & Friis, 1971), and my observations on endogenous env expression were independently confirmed by Hanafusa et al. (1970). Later, we detected alpha-retrovirus DNA in normal chick embryo cells by annealing methods (Varmus et al., 1972). Endogenous retroviral genomes were detected for the beta-retrovirus, murine mammary tumour virus in GR mice (Bentvelzen et al., 1970) and for the gamma-retrovirus, murine leukemia virus (MLV) in AKR mice (Lowy et al., 1971).

With the advent of Southern blotting, it became feasible to map integrated ERV loci. We showed that the ALV-related ERV genomes were already present in the chicken’s ancestor, the Red Jungle Fowl, but were absent from the three other extant species of the same genus, Gallus (Frisby et al., 1979). We concluded that this ERV had entered the chicken germ line within the past one million years, before domestication but after speciation, and we also found related genome sequences in other genera of Galliformes such as pheasants. Many further avian ERVs have since been detected (Bolisetty et al., 2012). Murine gamma-retrovirus ERVs have been thoroughly studied in inbred lines of mice (Nellåker et al., 2012). The complexity and the opportunities for recombination between ERVs is exemplified by tracing of the origin of a xenotropic MLV variant which spread to contaminated human prostate cancer specimens through xenografts and contamination of cultures and was mistakenly associated with chronic fatigue syndrome (Cingöz et al., 2012).

Whole genome sequencing has led us to appreciate that every vertebrate species examined contains multiple types of endogenous retrovirus (Magiorkinis et al., 2013; Stoye, 2012; Weiss & Stoye, 2013). Approximately 8% of human DNA is derived from fossil germ-line infection by retroviruses (Griffiths, 2001) and a much larger proportion of our genome is represented by retrotransposons such as LINE elements (Rebollo et al., 2012). I say ‘fossil infection’ because ERVs have the hallmarks of having been acquired by infection (Belshaw et al., 2004) rather than being relics of an earlier RNA world (Villereal, 2005). Furthermore, cDNA genome fragments of other types of RNA virus are also found in mammalian chromosomes (Feschotte & Gilbert, 2012). Retroviruses also integrate into the genomes of DNA viruses and there is evidence that they may have been spread iatrogenically through live attenuated vaccines for the avian DNA viruses of Marek’s disease and fowlpox (Niewiadomska & Gifford, 2013). An endogenous gamma-herpesvirus has recently been detected in the DNA of primitive primates (Aswad & Katzourakis, 2014), and integrated human herpesvirus 6 (HHV-6) is present the germ line of some humans (Pellett et al., 2012). Thus, the accumulation of viral sequences in the DNA of hosts and of other viruses is a broader phenomenon than ERVs alone.

**Evolutionary dynamics of exogenous and endogenous retrovirus**

There is a stark difference between the replication rate of exogenous, replicating viruses and that of endogenous viruses embedded in the host genome (Aiesakun & Katzourakis, 2015). The former undergo millions of replication cycles, including the reverse transcription step that is not subject to editing or repair of errors, whereas ERVs by definition are part of the host and only replicate at the pace of the host germ line. However, exogenous retroviruses may recombine at relatively high frequency with related ERV genomes if their transcripts are expressed as packageable RNA, converting elements of the ERV into an infectious virus again (Weiss et al., 1973). Many ERVs have become defective since entering the host germ line; for instance there are no known infectious ERVs in humans and dogs, although ERV genomes capable of being activated to infectious form are maintained in other species such as cats (Mata et al., 2015; Shimode et al., 2015) and pigs, which led us to examine the potential infection hazards of pig-to-human xenotransplantation (Le Tissier et al., 1997; Patience et al., 1997).

Host restriction factors may have evolved in part to protect against retroviral invasion (Sanz-Ramos & Stoye, 2013). On the other hand, ERVs often exhibit xenotropism, that is, when activated to produce infectious virus particles, these can more readily infect foreign host species. The Env glycoprotein of ERV can itself act as a restriction factor against exogenous infection by blocking receptors (Malfavon-Borja & Feschotte, 2015), a phenomenon that we first reported for avian ERV (Payne et al., 1971).
It is clear that replication-competent ERVs can make large leaps to infect distant host taxa; for instance, a baboon ERV crossed into cats to become a replication-competent ERV, RD114 (Benveniste & Todaro, 1974; Weiss, 2006). RD114 has a wide species host range in vitro and is known to re-infect cats in vivo (Shimode et al., 2015). Retroviral and lentiviral vectors pseudotyped by envelopes of RD114 and the gibbon virus (see below) are efficient in infecting human haematopoietic stem cells (Porter et al., 1996; Relander et al., 2005).

Another interesting example of cross-species infection and subsequent endogenization by replication-competent ERVs concerns gamma-retroviruses related to the gibbon ape leukemia virus (GALV). The first virus of this group to be isolated was a helper virus to a sarcoma virus in a pet woolly monkey that shared a home with a gibbon (Wolfe et al., 1971). Further isolates of GALV were made among captive lar gibbons (Hylobates lar) in Thailand (Kawakami et al., 1972, 1978), although this exogenous retrovirus has not been recorded in wild gibbons. The animals were used by the South-East Asia Treaty Organization for medical research, including inoculation with malaria, whole human blood and brain tissue from kuru autopsies (presumably from New Guinea), procedures that provided opportunities for cross-species infections. ERV sequences related to GALV were detected in the DNA of South-East Asian rodents such as Mus caroli (Lieber et al., 1975), Mus cervicolor (Benveniste et al., 1977), Vandeleuria oleracea (Callahan et al., 1979) and Melomys burtoni (Simmons et al., 2014).

In 1988, retroviral particles were reported in association with leukaemia in a koala by the late Daria Love (Canfield et al., 1988) who had earlier been my first post-doctoral fellow; (Love & Weiss, 1974). When the koala retrovirus (KoRV) was isolated and characterized, it was found to be closely related to GALV (Hanger et al., 2000). The koala is an Australian marsupial species (Phascolarctos cinereus) with no natural contact with gibbons or South-East Asian mice. Recent investigations of preserved taxidermy specimens indicate that KoRV has been present in koalas for at least 120 years (A´ vila-Arcos et al., 2015), but the epidemic of leukaemia seems to be more recent and has been spreading among koalas from high prevalence in the north-eastern regions of Australia to lower prevalence in the south (Tarlinton et al., 2006). The leukaemia appears to be related to an envelope variant (KoRV-B or KoRV-J) co-existing with the original form (KoRV-A) (Shojima et al., 2013; Xu et al., 2013), and it is not yet clear which variant is the main infectious virus. While KoRV-A utilizes the Pit-1 receptor like GALV and FeLV-B, KoRV-J utilizes thiamine transport protein 1 like FeLV-A (Shojima et al., 2013).

It thus appears that an ERV of rodents has crossed host species at least twice, to become an exogenous pathogenic retrovirus in both gibbons and koalas. What's more, KoRV is in the process of becoming endogenized in the koala germ line (Tarlinton et al., 2006; Ishida et al., 2015), just as RD114 became endogenized in cats after horizontal transfer from baboons (Benveniste & Todaro, 1974). Since many rodent species of the family Muridae carry related ERVs, we may be confident that the precursor of GALV evolved in rodents and spread secondarily to the koala and the ape (Fig. 2). The precise species giving rise to GALV and KoRV remains to be determined. Melomys burtoni lives in Northern Australia and Southern New Guinea and overlaps with the northern geographic range of koalas. However, since GALV and KoRV are more closely related to each other than to the ERV of M. burtoni (Simmons et al., 2014), it is unlikely that M. burtoni represents the immediate progenitor of KoRV.

Are endogenous retroviruses useful or harmful to the host?

ERVs and other retrotransposons are open to natural selection operating on the host. There has been much discussion about both deleterious and beneficial effects of ERVs (Moyes et al., 2007; Stoye, 2012) and new ERV insertions may disrupt essential host genes (Nell´aker et al., 2012). ERVs can be oncogenic, contributing to leukaemia in mice and cats (Lowy et al., 1971; Neil et al., 1991) and to mammary cancer in GR mice (Bentvelzen et al., 1978; Ross, 2010). Whether human ERV sequences (HERV) play a causative role in human cancers is less definitive (Bhardwaj & Coffin, 2014; Magiorkinis et al., 2013). Beta-retrovirus HERV-K (HML2) genomes have been implicated in testicular tumours (Löwer et al., 1996), melanoma (Schmitt et al., 2013) and breast cancer (Salmons Lawson & Günsburg, 2014), and specific HERV-K loci are more highly expressed in tumour tissue (Flockerzi et al., 2008). However, that does not yet prove HERV-K causality because one can reverse the logic to argue that tumour cells impose fewer restrictions on HERV-K expression.

HERV-K elements are a well-studied group of human ERVs and they are up-regulated by HIV infection (Vincendeau et al., 2015). There are few copies of ERV genomes related to HERV-K in New World primates but they underwent great expansion in Old World primates, and some full length human HERV-K genome integrations have been acquired since the divergence of humans and chimpanzees (Belshaw et al., 2004; Subramanian et al., 2011). HERV-K expression occurs in pre-implantation embryos; it is linked with pluripotency and becomes transcriptionally silenced upon differentiation (Fuchs et al., 2013; Grow et al., 2015). HERV-H and LINE elements are also tightly linked to transcriptional regulation in pluripotent cells and may exert an influence on early development (Pa´es et al., 2013; Robbez-Masson & Rowe, 2015).

One of the most striking examples of ERVs becoming a benefit to the host is the role of Env in driving cell fusion to form the syncytio-trophoblast of the mammalian placenta. Since we had employed cell fusion assays to study cell receptors for mammalian retroviruses (Nagy et al., 1983; Sommerfelt & Weiss, 1990), I became intrigued by the high expression in the human placenta of a defective
HERV with an open reading frame for env, ERV-3, (Boyd et al., 1993). We found that ERV-3 expression was tightly linked to trophoblast cell fusion (Fig. 3) and postulated that a functional retroviral Env glycoprotein would be able to effect the cell-to-cell fusion (Venables et al., 1995). The fusigenic property of ERV-3 was then demonstrated by Rote’s group upon expressing it in the choriocarcinoma cell line BeWo (Lin et al., 1999). However, our hypothesis appeared to be destroyed when Thierry Heidmann’s group showed that some humans lack the ERV-3 genome although they must have been born following a successful pregnancy with a presumably healthy placenta (de Parseval & Heidmann, 1998); indeed, gorillas also lack ERV-3 (Hervé et al., 2004). In fact, the hypothesis survived but with a different ERV Env glycoprotein. Blond et al. (2000) and Mi et al. (2000) showed that the Env of HERV-W (named syncytin) is expressed in the human syncytio-trophoblast and induces cell-to-cell fusion.

Placental HERV expression may have a dual role: to be locally immunosuppressive to protect the fetus from maternal rejection, and to induce the trophoblast to form syncyta (Boyd et al., 1993; Lavialle et al., 2013). The syncytin story has grown more intriguing with the demonstration that a second human ERV, HERV-FRD also encodes a distinct syncytin (Blaise et al., 2003) and that different orders of placental mammal employ Env glycoproteins of quite different ERVs to induce cell fusion of the trophoblast (Lavialle et al., 2013). Since one may assume that the evolution of the placenta was a monophyletic event, why have placental mammals repeatedly entrained different ERVs to effect trophoblast differentiation into a syncytium? Perhaps the panoply of ERVs resident in mammals offered continued opportunities to improve the cell fusion process in the wide variety of placenta among different mammalian taxa. And clinging on to our original hypothesis, perhaps ERV-3 was the default fusion inducer before higher primates adopted HERV-W and HERV-FRD, with some functional redundancy during the switch?

Transmissible tumour cells

Early in the era of AIDS, two notes were published speculating that canine transmissible venereal tumour (CTVT) may be a model for Kaposi’s sarcoma (KS) in AIDS (Hayes et al., 1983; Rechavi et al., 1991). Many veterinarians accepted that the transmissible factor was the tumour itself rather than an oncogenic virus (Cohen, 1985), but the oncologists and immunologists to whom I spoke were profoundly sceptical. The discovery of human herpesvirus 8
cells had spread among affected dogs worldwide from an ancient breed or possibly from grey wolves (Murgia et al., 2006). Further research led by Liz Murchison using next generation whole genome sequencing of CTVT specimens from different continents led us to revise the estimate of the time of its emergence to approximately 11 000 years ago in an ancient breed like the Malamute or Siberian Husky (Murchison et al., 2014). CTVT currently occurs in at least 90 countries, especially among stray animals (Strakova & Murchison, 2014). Although the tumour has clocked up thousands of somatic mutations and massive chromosomal rearrangements, it provides insight into the genome of an ancient dog; there is no need to dig in the permafrost for ancient canine DNA when an ancient tumour is still extant (Murchison et al., 2014; Parker & Ostrander, 2014).

We observed that one genetic marker in CTVT, a 722 bp sequence in the control region of mitochondrial DNA, was more diverse than expected and clustered into two distinct main clades. We therefore hypothesized that the tumour might have acquired host mtDNA during thousands of years of serial passage among dogs (Murgia et al., 2006). Sequencing and analysis of complete mtDNA sequences of diverse CTVT specimens demonstrated that the mtDNA is indeed polyphyletic (Rebbeck et al., 2011). Here we have an example of a malignant cell that has emerged as an infectious parasite in its own host species, and that host mitochondria have secondarily colonized the tumour cell, perhaps providing selective advantage to tumour cell growth.

CTVT is not unique and has served as a model of other ‘parasitic’ tumours (Weiss & Fassati, 2015). The best known is the devil facial tumour disease of the marsupial carnivore, the Tasmanian Devil (Sarcophilus harrisii). The Devil is currently threatened with extinction by this highly aggressive transplantable tumour (Murchison et al., 2012; Pearse & Swift, 2006) because this tumour is usually fatal, whereas dogs with CTVT often survive. Another example, just published, is a leukaemia of soft-shelled clams (Metzger et al., 2015) in which the tumour cells are thought to be dispersed in seawater and infection occurs through filter feeding (Weiss & Fassati, 2015).

The emergence of transmissible tumour cells remains a rare phenomenon and we do not fully understand how the tumours evade the host immune response. The lack of major histocompatibility antigens or their down regulation would promote the chance of tumour emergence (Murgia et al., 2006; Siddle & Kaufman, 2015). Looking into the literature, I found only one example of ‘naturally’ transmissible tumours, reported 50 years ago, among inbred strains of laboratory rodents; this was a histiocytic leukaemia in Syrian hamsters and remarkably, it could be transmitted by mosquitoes (Banfield et al., 1965).

In humans, there are numerous examples of occult tumour transmission from donors to immunosuppressed transplant patients (Laebznik & Parris, 2015). Leukaemia has

Fig. 3. Indirect immunofluorescence detection of expression of human ERV-3 envelope glycoprotein in the syncytiotrophoblast of a full term placenta. The intense staining of Env and the lack of cell membranes between nuclei is evident. Reproduced with permission from Venables et al. (1995). Bar, 20 μm.
also been transmitted *in utero* between fetuses sharing a placenta and from mother to child or *vice versa* (Greaves et al., 2003). It is a form of micro-chimerism like freemartins in cattle with placental anastomoses (Owen, 1945), but with greater pathological consequences. Among invertebrates, natural chimerism of non-malignant cells is a widespread phenomenon (Rinkevich, 2011). In colonial ascidians (*Tunicata*), both somatic cell invasion and germ cell migration are found where the germ cells of one genetic identity take up residence in the soma of another (Stoner & Weissman, 1996). Overall, there is plenty to speculate on the evolutionary dynamics of cells occupying another host’s body, which has been called ‘extracorporeal metastasis’ (Lazebnik & Parris, 2015).

### Conclusions

So what’s the microbe and what’s the host? Viruses can embed themselves in host chromosomes, and host genes can be transferred to viruses. Endogenous retroviruses can lie ‘dormant’ in the germ line for millions of years and re-emerge as replication-competent viruses infecting hosts of distant species. Endogenous retroviral genes can also be turned to good use by the host. Host cells can emerge as transmissible malignant clones, survive longer than any other vertebrate somatic cells and then be colonized in turn by mitochondria from new hosts many transplant generations later. There is enormous fluidity of genomes and cells within the eukaryotic world.

### Acknowledgements

I am most grateful to my post-doctoral mentors, Jan Svoboda in Czechoslovakia and Peter Vogt in USA for a stimulating intellectual environment, as well as the research students and post-doctoral members of my laboratory over many years. My research into oncogenic viruses was supported by Cancer Research UK and its precursor organizations, HIV research by the Medical Research Council, the Bill & Melinda Gates Foundation and the European Union, and investigations into transmissible tumour cells by the Wellcome Trust. I thank Rachael Tarlinton and Yasuhiro Takeuchi for constructive criticism of the manuscript.

### References


