2.1 Norovirus recombination

2.1.1 Definitions and general aspects of RNA recombination

The extremely high genetic variability observed in RNA viruses can be attributed to three particular evolutionary forces: point mutation, reassortment (exclusively undergone by segmented RNA genomes) and recombination (Domingo and Holland, 1997). The high rate of nucleotide substitutions, due to short replication times, high multiplicity and a lack of proofreading activity of RdRp, allows RNA viruses to continuously generate mutant genomes (Holmes, 2010). Causing more drastic evolution jumps than point mutations, RNA recombination involves the exchange of genetic information between non segmented RNA genomes. However, genetic divergence is restricted by the necessity to maintain a functional viral RNA genome and by environmental selective pressures. Therefore, RNA recombination not only enables the creation and spread of advantageous traits but could also rescue viral genomes by repairing mutation errors in essential viral genes (Lai, 1992).

History

Up to the 60's, recombination was confined exclusively to DNA molecules and sequence exchange between RNA molecules was first discovered *in vitro* for poliovirus - member of *Picornaviridae* family - in 1962 (Hirst, 1962). By the use of polioviruses possessing genetic markers such as guanidine resistance and resistance to horse serum, recombinants exhibiting the resistant phenotypes of both parental viruses were isolated at 15-20 times higher frequency during mixed cell culture infections than during single infections. This was the first evidence of RNA recombination in non segmented genomes and the phenomenon has been described for an increasing number of RNA viruses since (Aaziz and Tepfer, 1999).

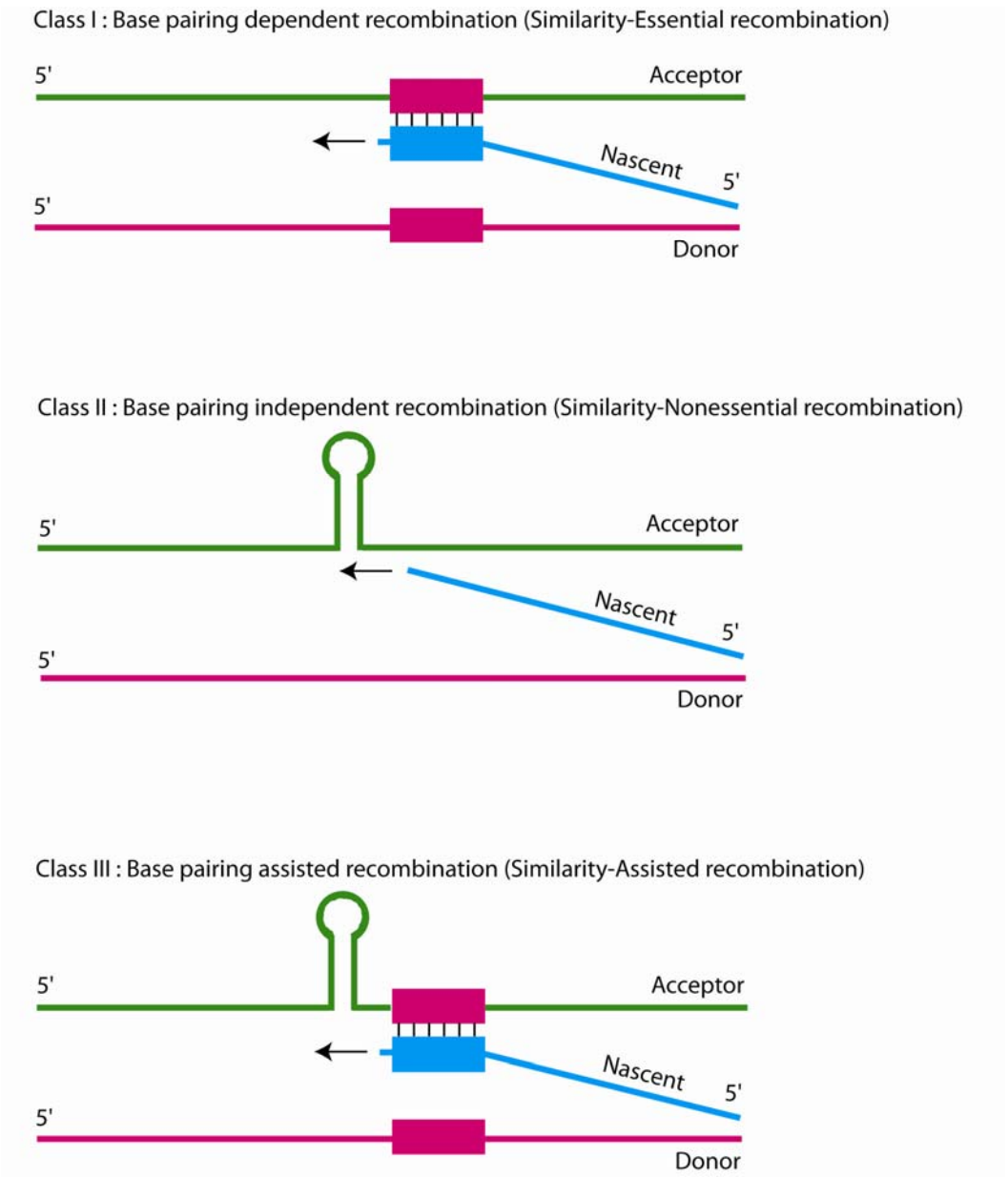

Figure 11: Types of RNA recombination. Replicase-mediated RNA synthesis after the template-switch events is shown by an arrow. The hairpin structure shown symbolically represents various RNA features that are required for Class II and Class III recombination (Figure adapted from Nagy and Simon, 1997).

Types of RNA recombination

Based upon sequence similarities between the 'donor' and 'acceptor' parental RNA sequences and other RNA requirements at the site of crossing over, replicative RNA recombination of model viruses characterised to date can be categorised into three types or classes: similarity-essential (class 1, with either precise or aberrant end products), similarity non-essential (class 2) and similarity-assisted (class 3) (Nagy and Simon, 1997) (Figure 11). **Base pairing dependent recombination (similarity-essential recombination)** requires substantial sequence similarity between parental RNAs, and encloses both the homologous and aberrant homologous recombination events described by Lai in 1992 (Lai, 1992). A majority of recombination events between full-length viral genomes involve homologous parental RNAs and crossovers at homologous sites that result in a neat replacement of the acceptor sequence by the donor sequence leaving the structure unchanged. In the absence of a strict alignment of the parental RNAs, recombination crossovers may occur and lead to sequence insertions or deletions. **Base pairing independent recombination (similarity-non essential recombination)** does not require sequence similarity but may be present. In this model recombination may be determined by other RNA features, such as RNA polymerase binding sites (highly structured RNA), secondary structures, and heteroduplex formation between parental RNAs. Finally, **base pairing assisted recombination (similarity-assisted recombination)** combines features of class I and class II recombination. Sequence similarity influences sites and frequency of recombination; however additional RNA features are also required. Hitherto, nearly all recombination studies in RNA viruses have supported a copy-choice model (Cheng and Nagy, 2003; Jarvis and Kirkegaard, 1992; Kirkegaard and Baltimore, 1986; Nagy and Simon, 1997). Under this model, hybrid RNA molecules are the result of the switching of RdRp complex from one RNA molecule to another during replication. This template switching mechanism is completely different from the enzyme-driven breakage-rejoining mechanism of recombination in DNA that is not as intimately linked to replication. Subsequently a non-replicative recombination model has been introduced for polioviruses but this model still needs to be confirmed for other RNA viruses (Gmyl *et al.*, 1999).

Recombination detection tools

Evidence of recombination can be provided by two different approaches. The first approach is based on direct laboratory experimental data obtained by making use of non-replicative variant viruses that recover functionally through recombination. These experimental studies are conducted under strong selective pressure conditions and have allowed the isolation of recombinants for numerous animal and plant RNA viruses (Alejska *et al.*, 2005; Hirst, 1962; Lai *et al.*, 1985; McCahon and Slade, 1981). The second approach is based on phylogenetic analysis of RNA viral genomes. Indeed, the development of PCR constituted a revolution in the study of virus recombination both in natural conditions and in the laboratory. The occurrence of recombination is suggested by discordant genetic relatedness when different parts of the genome are considered. A multitude of graphical tools either used sole or in combination with each other, are available and offer the advantage to recover ancient or exceedingly rare recombination events. Moreover, they provide information on the precise location of the putative crossover points.

Evolutionary advantages of RNA recombination

Hypothesises on the biological signification of RNA recombination have evolved from unavoidable side effect of the RdRp during replication to indispensable source of genetic diversity (Domingo and Holland, 1997). Recombination enables the exploration of new genetic combinations leading to i) enhanced viral fitness (the adaptation of an organism or a virus to a given environment), ii) the rescue of fit viral genomes from debilitated parental genomes, or iii) new starting points for subsequent viral evolution. Indeed, 18% of the infections in the global HIV-1 pandemic are circulating recombinant forms of HIV-1 viruses that might dispose of a selective advantage upon other HIV viruses (Buonaguro *et al.*, 2007). Thus, recombination, by contributing to the genetic breadth and viral population plasticity that characterises RNA viruses, may be responsible of phenotypic switches that give rise to new virus strains with unpredictable virulence. As a matter of fact, a great majority of the emergent or re-emergent human viral pathogens during the last decades were associated with RNA viruses that display active recombination or reassortment (Domingo, 2010). For example, multiple reports in literature implicate homologous recombination and host shifting

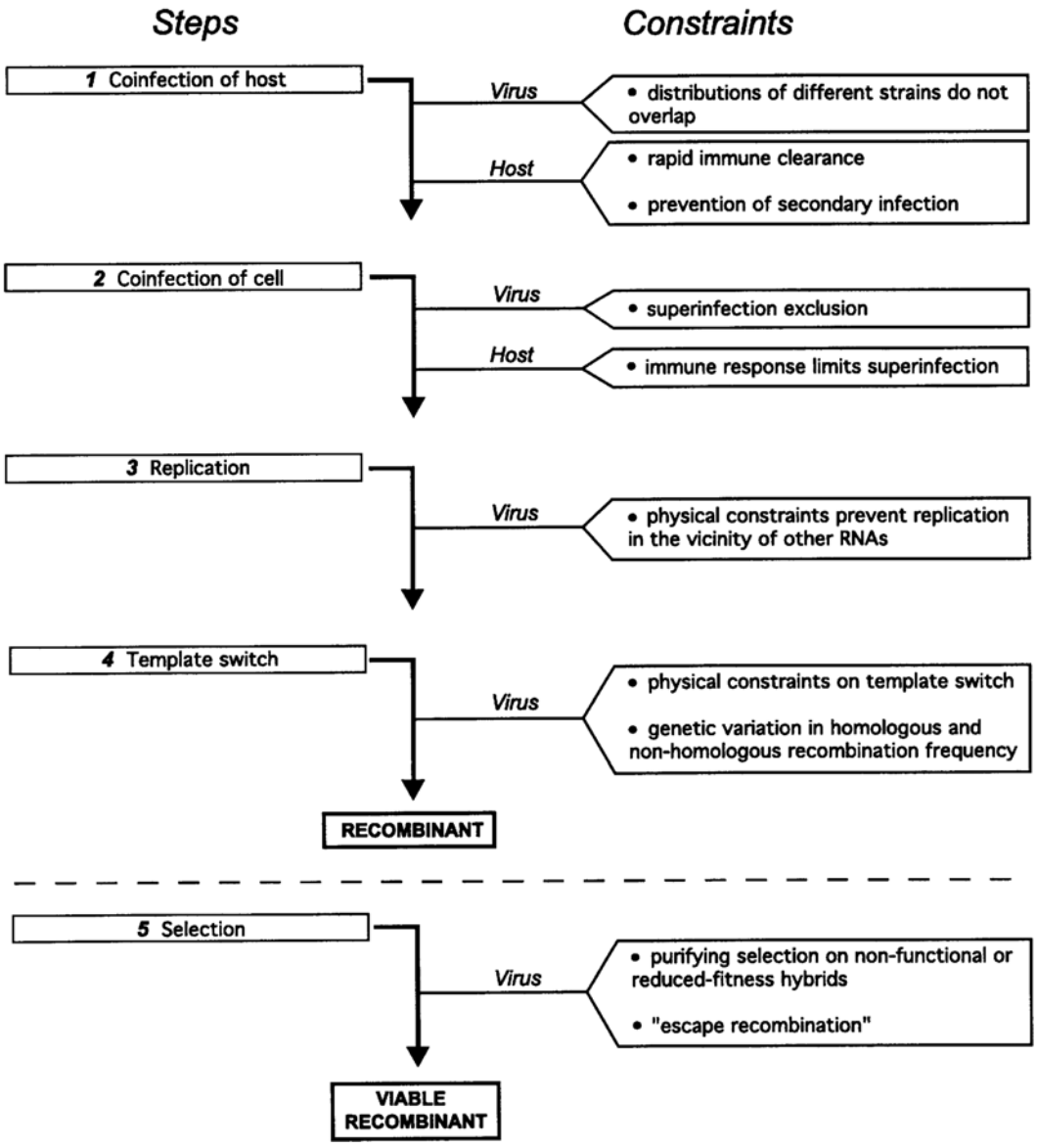

Figure 12: Model of RNA recombination by copy-choice proposed by Worobey and Holmes (1999). Requirements for the production of viable recombinant viruses are divided into 5 steps. Virus or host constraints could block recombination by limiting or preventing the completion of any of the five steps.

in the phylogenetic history of SARS-coronavirus (Graham and Baric, 2010). Also, several recent poliomyelitis outbreaks have been shown to be involved with viruses generated by recombination between attenuated vaccine polioviruses and other circulating enteroviruses (Simmonds, 2006).

Constraints on RNA recombination

Evidence of recombination has been found for a continuously growing list of viruses in natural or experimental conditions at different frequencies indicating not all RNA viruses are equally prone to recombination. Different prerequisites need to be fulfilled in order to obtain successful viral RNA-RNA recombinants (Worobey and Holmes, 1999). First, an individual must be simultaneously infected by two different virus strains. Then, the divergent viruses need to coinfect a single cell and replicate in close intimacy if recombination is to proceed. Finally, recombination takes place when the viral replicase namely RdRp actually switches templates resulting in the creation of a hybrid RNA molecule between the two parental genomes. Non-recombining viruses are the result of any constraint impeding one of the above mentioned requirements preventing RNA recombination from taking place. These factors are not only of viral origin, environmental and host factors have been shown to play a key role on the emergence of viral RNA recombinants (Jaag and Nagy, 2010). A final fifth step (non essential for recombination) was described; the removal by purifying selection of non-replicative or reduced-fitness hybrids implies that only viable or highly fit recombinants will be detected. All steps leading to RNA recombination including possible constraints are schematised in figure 12.

2.1.2 Recombination in noroviruses

Like all RNA viruses, NoVs show high mutation rates due to the lack of proofreading repair mechanisms of the RdRp. Mutation rates for NoV have been estimated to range between 2-9 10^{-3} substitutions per nucleotide per year (Bok *et al.*, 2009; Bull *et al.*, 2010; Siebenga *et al.*, 2010b). NoV strain diversity is not exclusively due to nucleotide substitutions, recombination was also shown to occur. The first description of naturally occurring recombination dates from 1997 when incongruent nucleotide sequence similarities were found for prototype Snow

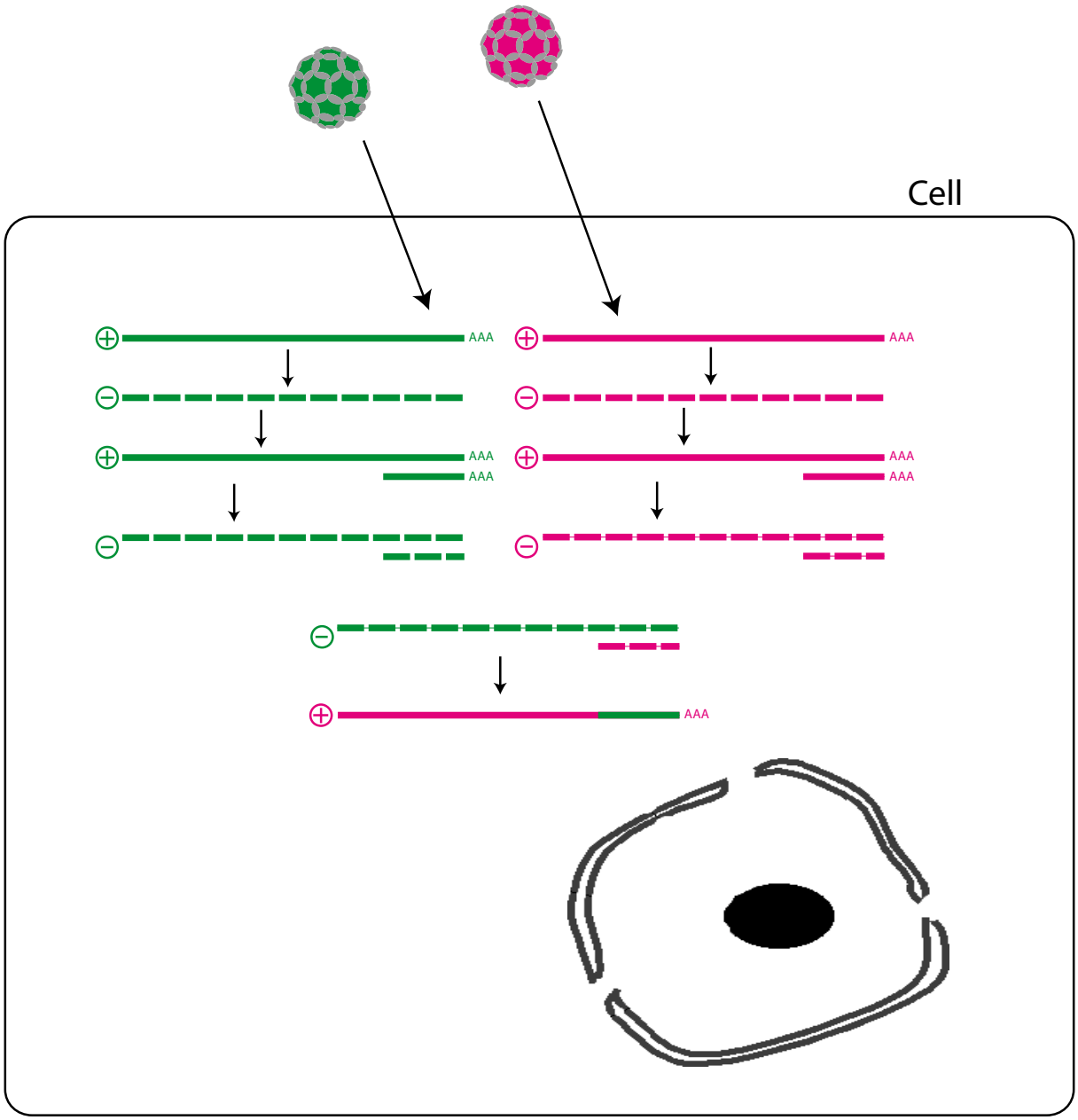

Figure 13: Mechanism for recombination in norovirus as proposed by Bull et al. (2005). Recombination occurs when the RNA dependant RNA polymerase switches templates from the full-length negative strand from one parental virus to an available negative subgenomic RNA from another parental virus.

Mountain Virus in the capsid and RdRp region (Hardy *et al.*, 1997). The increased awareness for NoV recombination following its discovery resulted in increased reporting of NoV recombinants in natural conditions across the globe (Ambert-Balay *et al.*, 2005; Jiang *et al.*, 1999a; Phan *et al.*, 2006; Reuter *et al.*, 2006). This far, at least 22 NoV recombinant types have been described within GI (1), GII (16), GIII (2), GIV (1) and GV (2) along with one intergenogroup recombinant GI/GII (Bull and White, 2010). In most cases, crossovers were located within or close to the ORF1/ORF2 overlap suggesting the start of the caspid gene constitutes a recombination hotspot for NoVs. A simple model, according to the similarity-assisted recombination model, has been proposed for NoV recombination based upon RNA replication and the synthesis of sgRNA by internal initiation (Bull *et al.*, 2005) (Figure 13). This model was divided in four subsequent steps. First, negative full-length genome intermediates (dashed lines) are transcribed by the RdRp. Second, positive stranded genomic and sgRNA (straight lines) are generated after RdRp binds to RNA promoter sequences located at the start of ORF1 and ORF2. Third, negative full-length gRNA and sgRNA species are synthesised from positive RNA. Finally, recombination occurs by template switching of the RdRp during positive strand synthesis. RdRp stalls due to a complex secondary structure at the subgenomic promoter before it switches templates to an available negative (or full-length genome) sgRNA species. Indeed, this region has been shown to exhibit a marked suppression of synonymous variability that coincides precisely with stem-loop RNA secondary structures on the anti-genomic strand upstream of a subgenomic transcript within each genus of the family *Caliciviridae*, including NoV (Simmonds *et al.*, 2008). In the absence of a cell culture system, NoV recombination has been exclusively studied by phylogenetic analysis but the discovery of a valuable murine model has opened perspectives in the study of NoV recombination. Predictive recombination tools together with similarity plots between putative recombinant genomes and the suspected parental genomes have suggested recombination at breakpoints within ORF2 in several MNV genomes (Thackray *et al.*, 2007), but experimental *in vitro* or *in vivo* evidence of RNA recombination between co-infecting NoV strains is still lacking.