Proposal from *Circoviridae* study group :

1. Proposals
2002.V026.02: Establish a new genus of vertebrate viruses with circular, single-stranded DNA
2002.V027.02: Name the novel genus *Anellovirus*
2002.V028.02: Designate *Torque teno virus* (TTV) as the type species of the genus

2002.V029.02. designate *Torque teno mini virus* (TTMV) as a species member of the newly created genus *Anellovirus*

2. Purpose : The purpose of this proposal is to initiate the taxonomic classification of recently-recognised viruses, known as "TT virus" and "TT virus-like mini virus". TT virus and TT virus-like mini virus were identified in 1997 and 1999, respectively. They contain circular single-stranded DNA genomes of about 3800 nts (TT virus) or 2900 nts (TT virus-like mini virus). As such, the issue of their taxonomy has been addressed by the *Circoviridae* Study Group.

2.1. General characteristics:

a) Genomic organisation and expression TT virus. At the present time, the 61 full or nearly full-length sequences deposited in databases originate from various countries (Africa, China, USA, France, Ghana, Indonesia, Japan and United Kingdom). Sequence analysis revealed a similar genomic organisation with the presence of a coding region of about 2600 nts-long and a non-coding region of about 1200 nts-long. The latter harbours a region of about 110 nts-long with high GC content (~90%) (Okamoto et al., 2001a; Peng et al., 2002). Coding region: Two main open reading frames, ORF1 and ORF2, may be deduced directly from the nucleotide sequence. These ORFs partially overlap and their estimated sizes slightly differ between isolates. ORF1 is composed of about 760-770 amino acids (aa), and may encode a structural protein corresponding to the capsid of the virus. Moreover, the presence of conserved aa motifs, which occur in the replication-associated (Rep) proteins of other animal and plant viruses with circular single-stranded DNA genomes, suggests that replication of TT virus DNA uses a rolling circle replication (RCR) mechanism. Current data suggest that ORF2 may encode about 120 aa, which probably constitutes a non-structural protein possibly involved in viral replication. Non coding region: This GC-rich region represents about 30% of the genome and contains several regulatory motifs. mRNAs: At least 3 mRNAs of different sizes (2.9 kb, 1.2 kb and 1kb) are transcribed from the negative strand of putative circular double-stranded replicative form (RF) DNA. The presence of these mRNAs not only supports the view that both ORF1 and ORF2 are functional, but also imply that 2 new ORFs, i.e. ORF3 (~260 aa) and ORF4 (~250 aa) can be obtained by complex splicing. The detection of viral mRNAs appears not to be restricted to a unique tissue but rather to various tissues and organs (Okamoto et al., 2001b).
b) TT virus-like mini virus Twelve complete sequences belonging to French and Japanese isolates are currently described. Genome sizes range from 2850 to 2950 nts. The genomic organisations appear to be conserved between isolates and coding (~2150 nts) and non-coding (~720 nts) regions might be deduced, the latter showing a GC-rich zone (~80 nts). The variability profile constructed with full-length sequences shows a unique zone, upstream from the coding region, which is highly conserved between isolates

(Biagini et al., 2001; Takahashi et al., 2000). Coding region: Like TT virus, the ORF profile corresponds to 2 main overlapping ORFs, ORF1 and ORF2. ORF1 is composed of about 675 aa and contains some conserved motifs characteristic of Rep proteins. ORF2 is composed of about 100 aa. Non coding region: It accounts for about 25% of the viral genome. Some conserved motifs are detected along with a sequence (GGGGGCTCCGCCCC) found in triplicate inside the GC-rich zone. mRNAs: The fact that TT virus and TT virus-like mini virus share similar transcription profiles is highly suggestive that several mRNAs may be expressed as for TT virus. This hypothesis would also confirm the functionality of ORF3 and ORF4.

c) Genetic variability: TT virus
TT virus harbours a high genetic heterogeneity for a DNA virus, as values above 50% for genetic variability may be observed between 2 complete genomic sequences. Moreover, the existence of 3 hyper variable regions was described in the centre of ORF1 (Bendinelli et al., 2001). The growing number of divergent TT virus isolates deposited in databases can be assigned to more than 25 genotypes of the virus at the present time. Genotype repartition appears not to be related to geographic origin.
TT virus-like mini virus. The analysis of complete or partial sequences available to date reveals a high genetic variability, comparable to that found for TT virus. Based on complete or partial (ORF1) nucleotide sequence analysis, 3 genotypes may be identified at the present time with a cut-off value of 40% for genetic distance. Moreover, the analysis of isolates belonging to distinct continents clearly demonstrated the absence of geographic cluster (Biagini et al., 2001).

d) Epidemiology: TT virus: The large number of epidemiological studies permitted to clearly demonstrate the global distribution of the virus (Africa, North and South America, Asia, Europe, Oceania), in rural and urban populations (Prescott et al., 1998). The TT virus prevalence in the general population is high (> 75%). Despite unproved link between TT virus infection and a given pathology, the hypothesis of a relation between viral load and the immune status of the host was suggested (Shibayama et al., 2001). Moreover, although initially suspected of being transmitted only by blood transfusion, the global dispersion of the virus in populations and its detection in various biologic samples (plasma, saliva, feces…) suggest combined modes of diffusion, and in particular the spread by saliva droplets. Other modes of transmission are also suggested by some studies, like maternal or sexual transmission of the virus. TT virus-like mini virus: TT virus-like mini virus prevalence studies revealed high values in the general population (~75%). As for TT virus, TT virus-like mini virus DNA appears detectable in various biologic samples (eg plasma, PBMCs, saliva, feces, cervical swabs): this suggests comparable routes of transmission for both viruses. Moreover, a recent study suggests an immune control of TT virus-like mini virus viral load (Gallian et al., 2002).

e) TT virus/TT virus-like mini virus and non-human hosts: Viruses were already detected in non-human primates (chimpanzee, macaque, tamarin, douroucouli) and in farm animals. The analysis of 5 complete non-human TT virus sequences reveals a high heterogeneity on the size of the viral genome (3371 to 3798 nts), along with a high genetic divergence for 4 out of the 5 sequences when compared with human isolates (Okamoto et al., 2000). However, genomic organisation and transcription profiles correspond to those found in human isolates.

2.2. TT virus/TT virus -like mini virus comparisons: TT virus and TT virus-like mini virus share comparable genomic organizations with 2 main overlapping ORFS (ORF1 and 2), presumed transcription profile, and conserved areas including a GC-rich region and a short region located upstream the ORF2. However, they differ in genomic size, and comparison of complete nucleotide sequences reveals a high genetic divergence between TT virus and TT virus-like mini virus isolates.

2.3. Comparison with other circoviruses: The *Circoviridae* family presently comprises 3 virus species: *Chicken anaemia virus* (CAV) is the only virus in the genus *Gyrovirus*; *Porcine circovirus* (PCV) and *Beak and feather disease virus* (BFDV) are classified as members of the genus *Circovirus* (Mankertz et al., 2000; Niagro et al., 1998; Todd et al., 2001). The comparison of the already known characteristics of these viral species with those concerning TT virus and TT virus-like mini virus highlights some major differences regarding: 1- The global distribution of TT virus and TT virus-like mini virus in human populations and in some non-human hosts 2- The very high prevalences found, at least, in human populations 3- The non association with a given pathology (at present time) 4- The possible co-infection of hosts with multiple strains, even highly divergent 5- The very high genetic variability 6- The considerable differences in the nucleotide sequence and genomic size 7- The specific genomic organisation 8- The specific transcription profile. Although as shown immediately above, TT virus and TT virus-like mini virus differ in many important ways from the classified circoviruses, the relationship of TT virus and TT virus-like mini virus to CAV is of particular interest to this proposal since, in terms of their molecular biology, they resemble CAV in a number of characteristics, namely:

1. All possess negative sense genomes, in contrast to the ambisense genomes possessed by circoviruses, that belong to the genus *Circovirus*.

2. Two of the 3 ORFs expressed by CAV have features in common with ORF1 and ORF2 specified by TT virus and TT virus-like mini virus. Thus, the largest CAV ORF, found closest to the 3` end of the coding region, encodes the capsid protein and, in this respect, it resembles the putative capsid proteins of TT virus and TT virus-like mini virus.

3. The capsid protein of CAV possesses aa motifs that are characteristic of RCR Rep proteins and, in this respect, it resembles the putative capsid proteins of TT virus and TT virus-like mini virus.

4. The ORF2 of CAV, found closest to the 5` start of the coding region, resembles the similarly- located ORF2s of TT virus and TT virus-like mini virus in that they all contain aa sequences that are characteristic of protein tyrosine phosphatases (PTPase). In the case of CAV, PTPase activity has been demonstrated in vitro using the recombinant DNA derived ORF2 protein Peters et al. (2001). If it can be shown that TT virus and TT virus-like mini virus also utilise a PTPase, a highly similar viral replication strategy is suggested for these viruses and CAV.

5. The non-coding regions of CAV is G:C rich and exhibits some nucleotide homology with TT virus and TT virus-like mini virus. However, CAV differs from TT virus

and TT virus-like mini virus in that it uses a single, unspliced polycistronic transcript, whereas with TT virus spliced transcripts have been detected and this is likely to be case with TT virus-like mini virus.

3. Revised organisation: The taxonomic implication is that TT virus and TT virus-like mini virus will be recognised as members of a new "floating" genus, named "*ANELLOVIRUS*", which is unattached to the family *Circoviridae*. Although it is recognised that these viruses share common features, relating to genome organisation and expression, with *Chicken anaemia virus*, at present we consider it appropriate to establish a new genus. This does not preclude the possibility of establishing, at some time in the future, a separate virus family for the genus *Anellovirus* to which the genus *Gyrovirus*, containing CAV might also be assigned.

4. Derivation of proposed names: The name of the genus "*Anellovirus*" is derived from "Anello", the ring, and relates to the circular nature of the DNA genome. A new name for TT virus is proposed as "Torque Teno Virus" (TTV). This name is derived from "Torques", the necklace and "Tenuis", thin, and relates to the circular, single-stranded nature of its DNA genome. Importantly, the acronym "TTV" is maintained, as this is the name currently used by those in the field and already present in more than 380 publications. A new name for TT virus-like mini virus is proposed as "Torque Teno Mini Virus" (TTMV). The name is related to TTV, as TTMV present some major characteristics retrieved in TTV (see item 2.2.).

References:

Bendinelli, M., Pistello, M., Maggi, F., Fornai, C., Freer, G. & Vatteroni, M. L. (2001). Molecular properties, biology, and clinical implications of TT virus, a recently identified widespread infectious agent of humans. Clin Microbiol Rev 14, 98-113.

Biagini, P., Gallian, P., Attoui, H., Touinssi, M., Cantaloube, J., de Micco, P. & de Lamballerie, X. (2001). Genetic analysis of full-length genomes and subgenomic sequences of TT virus-like mini virus human isolates. J Gen Virol 82, 379-83.

Gallian P., Biagini P., Attoui H., Cantaloube J.F., Dussol B., Berland Y., de Micco P. and de Lamballerie X. (2002). High genetic diversity revealed by the study of TLMV infection in French haemodialysis patients. Journal of Medical Virology In press.

Mankertz, A., Hattermann, K., Ehlers, B. & Soike, D. (2000). Cloning and sequencing of columbid circovirus (CoCV), a new circovirus from pigeons. Arch Virol 145, 2469-79.

Niagro, F. D., Forsthoefel, A. N., Lawther, R. P., Kamalanathan, L., Ritchie, B. W., Latimer, K. S. & Lukert, P. D. (1998). Beak and feather disease virus and porcine circovirus genomes: intermediates between the geminiviruses and plant circoviruses. Arch Virol 143, 1723-44.

Okamoto, H., Fukuda, M., Tawara, A., Nishizawa, T., Itoh, Y., Hayasaka, I., Tsuda, F., Tanaka, T., Miyakawa, Y. & Mayumi, M. (2000). Species-specific TT viruses and cross-species infection in nonhuman primates. J Virol 74, 1132-9.

Okamoto, H. & Mayumi, M. (2001a). TT virus: virological and genomic characteristics and disease associations. J Gastroenterol 36, 519-29.

Okamoto, H., Nishizawa, T., Takahashi, M., Asabe, S., Tsuda, F. & Yoshikawa, A. (2001b). Heterogeneous distribution of tt virus of distinct genotypes in multiple tissues from infected humans. Virology 288, 358-68.

Peng, Y.H., Nishizawa, T., Takahashi, M., Ishikawa, T., Yoshikawa, A. & Okamoto, H. (2002) Analysis of the entire genomes of thirteen TT virus variants classifiable into the fourth and fifth genetic groups, isolated from viremic infants. Arch Virol 147, 21-41.

Prescott, L. E. & Simmonds, P. (1998). Global distribution of transfusion-transmitted virus [letter]. N Engl J Med 339, 776-7.

Shibayama, T., Masuda, G., Ajisawa, A., Takahashi, M., Nishizawa, T., Tsuda, F. & Okamoto, H. (2001). Inverse relationship between the titre of TT virus DNA and the CD4 cell count in patients infected with HIV. Aids 15, 563-570.

Takahashi, K., Iwasa, Y., Hijikata, M. & Mishiro, S. (2000). Identification of a new human DNA virus (TTV-like mini virus, TLMV) intermediately related to TT virus and chicken anemia virus. Arch Virol 145, 979-93.

Todd, D., Weston, J.H., Soike, D. & Smyth, J.A. (2001). Genome sequence determinations and analyses of novel circoviruses from goose and pigeon. Virology 286, 354-62.

Peters, M.A., Crabb, B.S. & Browning, G.F. (2001) Chicken anaemia virus VP2 is a novel protein tyrosdine phosphatase. Proceedings of the Second International Symposium on infectious bursal disease and chicken infectious anaemia. Rauischholzhausen, Germany, June 2001.