The novel mitochondrial genome architecture of the potato cyst nematode *Globodera pallida*

M.R. Armstrong, V.C. Blok, M.S. Phillips & D.L. Trudgill

Mitochondria are oval-shaped organelles present in the cytoplasm of all eukaryotic cells. They are the site of the essential metabolic process of oxidative phosphorylation. Each mitochondrion contains a small chromosome which encodes products necessary for mitochondrial function. Typically, metazoan mitochondrial DNAs (mtDNAs) are single circular molecules between 13.5 and 30 kb in length that encode 22 transfer RNAs (tRNAs), two ribosomal RNAs (rRNAs) and 12 or 13 proteins involved in oxidative phosphorylation and electron transport. The inability to detect a mtDNA of sufficient length to encode the typical metazoan mitochondrial gene complement, in populations of the Potato Cyst Nematode, *Globodera pallida*, is a unique departure from conventional models of metazoan mtDNA structure. In *G. pallida*, at least six small circular mtDNA (scmtDNA) molecules are typically found, in populations from a wide variety of locations. These scmtDNA molecules vary in size from 6.3 to 9.5 kb (Fig. 1). No evidence for a larger 'wild type' molecule has been found using PCR, Southern blotting or electron microscopy (Fig. 2), although the presence of such a molecule at low levels

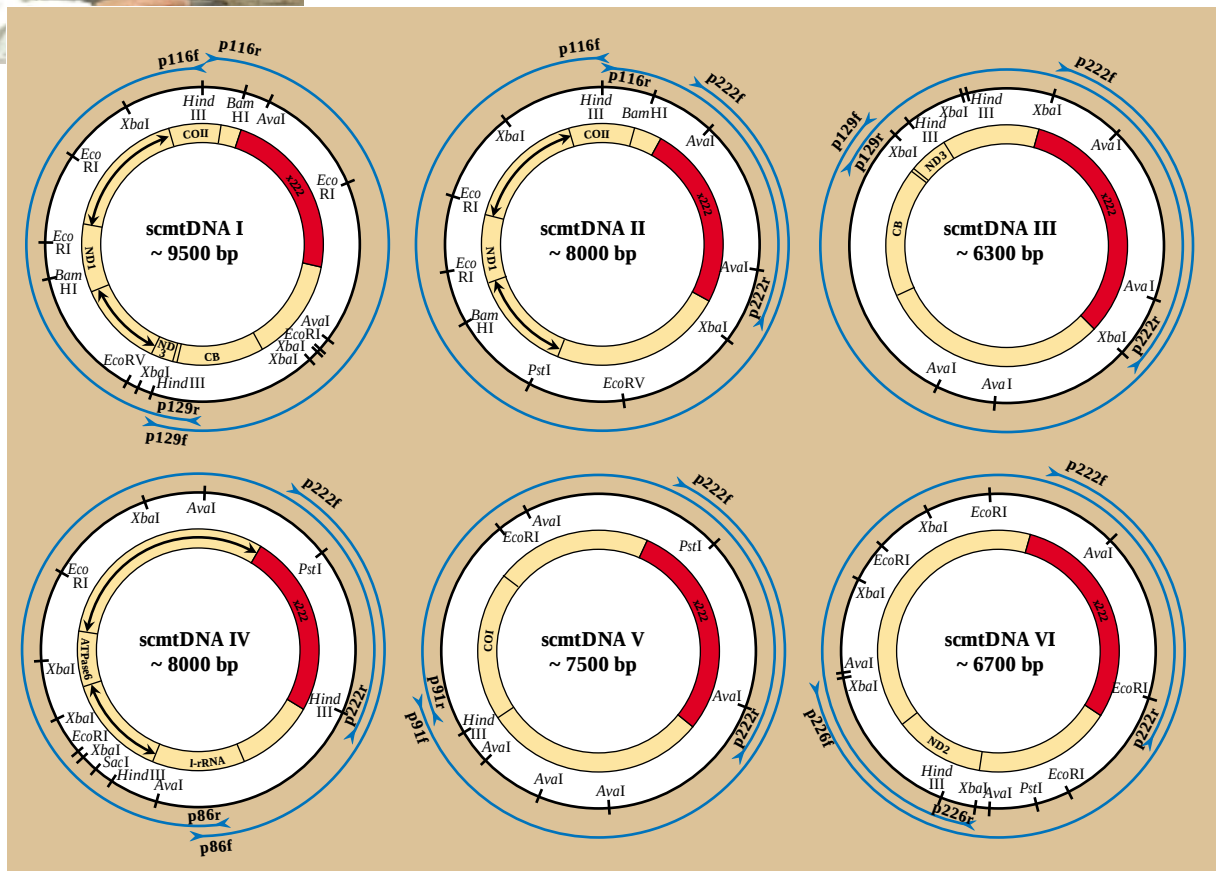

Figure 1 Restriction maps of the six scmtDNAs referred to in the text. The common ~2 kb non-coding region present on each scmtDNA is coloured red. Preliminary data from hybridisation studies regarding the distribution of mitochondrial gene coding sequences among the scmtDNAs is indicated. Note that scmtDNA IV alone contains rRNA sequences.

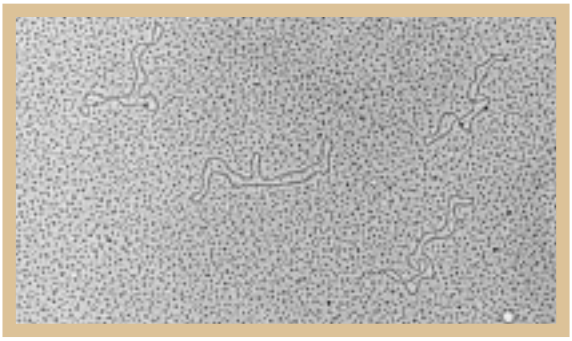

Figure 2 Electron micrograph of *G. pallida* scmtDNA molecules. The four molecules shown were estimated to range in size from 7.4 - 10 kb. No molecule longer than 10.5 kb was identified from over 150 such measurements.

cannot be ruled out. Each scmtDNA has been demonstrated to contain mitochondrial gene coding sequences and a ~ 2 kb non-coding region, which is common to all scmtDNAs. Preliminary investigations have suggested that all 12 mitochondrial protein coding genes found on nematode mtDNAs are distributed amongst the aforementioned scmtDNAs. In particular, rRNA genes (which are necessary for the translation of mitochondrial proteins) are found on only one scmtDNA. This raises the possibility that the components of the mtDNA operate in concert to encode the products necessary for mitochondrial function.

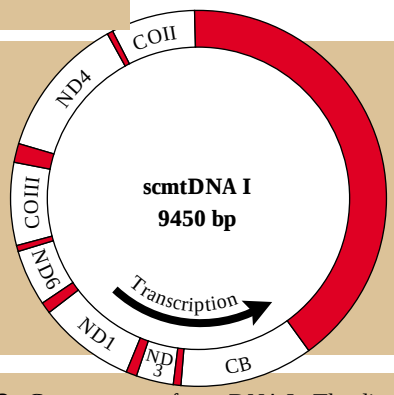

Figure 3 Gene content of scmtDNA I. The direction of transcription of the seven mitochondrial genes is indicated by the arrow. The 3545 bp non-coding region, containing a sequence similar to the non-coding regions present on all other scmtDNAs, is coloured red as are the six short intergenic sequences.

The complete sequence of scmtDNA I has been obtained. Analysis of this sequence indicates that it encodes seven full length mitochondrial proteins. Other than these genes, no regions of this molecule were identified that could encode further mitochondrial genes. Significantly, no rRNA genes were identified. These observations provide support for the notion that scmtDNAs in general are functional, while also confirming that scmtDNA I in particular would not be functional in isolation. The gene content of scmtDNA I is presented in Figure 3.

The presence of certain scmtDNAs was found to be unvarying in European and South American populations. For example, scmtDNA IV is invariably found. This might be expected given the presence of rRNAs on this scmtDNA. However, some scmtDNAs were found to vary enormously in their abundance . For example, in the majority of European populations, scmtDNAs II and III are the dominant mitotypes, while scmtDNA I is not detectable by Southern hybridisation. However, in an unusually virulent population (Luffness), scmtDNA I is the dominant mitotype ,while scmtDNAs II and III are undetectable (Fig. 4). This may suggest that the multipartite state affords a degree of flexibility in genome organisation as scmtDNAs I, II and III each contain some coding sequences in common (Fig. 1).

Potato cyst nematodes, particularly *G. pallida* in the U.K., are important agricultural pests. European populations exhibit considerable variation in virulence, but the lack of qualitative sources of resistance has prevented the establishment of useful phenotypic classifications among the majority of populations. Analysis of scmtDNA sequence variation has provided new insights as to the relationship between European *G. pallida* populations and their indigenous South American

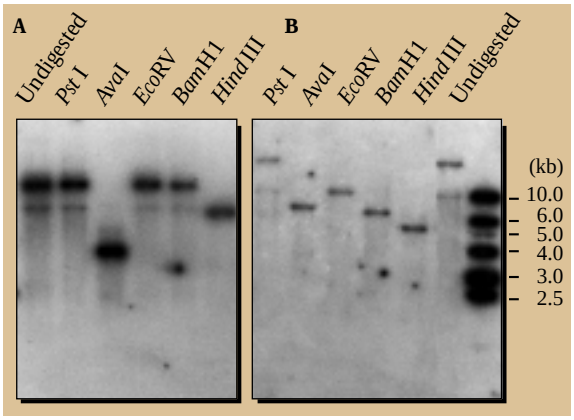

Figure 4 Southern blot of total genomic DNA from population Gourdie (Panel A) and Luffness (Panel B) hybridised with a probe specific to both scmtDNAs I and III. The restriction fragments detected in Panel A (Gourdie) are consistent with having originated from scmtDNA III. The restriction fragments detected in Panel B (Luffness) are consistent with having originated from scmtDNA I.

ancestors. Figure 5 shows a schematic representation of sequence variants of scmtDNAs I-IV found among five key populations, two South American and three European. South American populations contain a more diverse population of scmtDNA sequence variants than European populations. For example, only one of the four scmtDNAs typically found in European populations was detected in South American population P5A and this was a highly diverged sequence variant of scmtDNA IV. This suggests that this population has historically experienced low levels of gene flow with populations like P4A. Figure 5 suggests that population P4A also contains a diverse range of scmtDNA molecules. All the scmtDNAs found in the British populations Gourdie and Luffness are subsets of those present in population P4A. A third British population, Pa1, contains some novel scmtDNA sequence variants, suggesting it may have been derived from a distinct gene pool to that which gave rise to populations like Gourdie and Luffness.

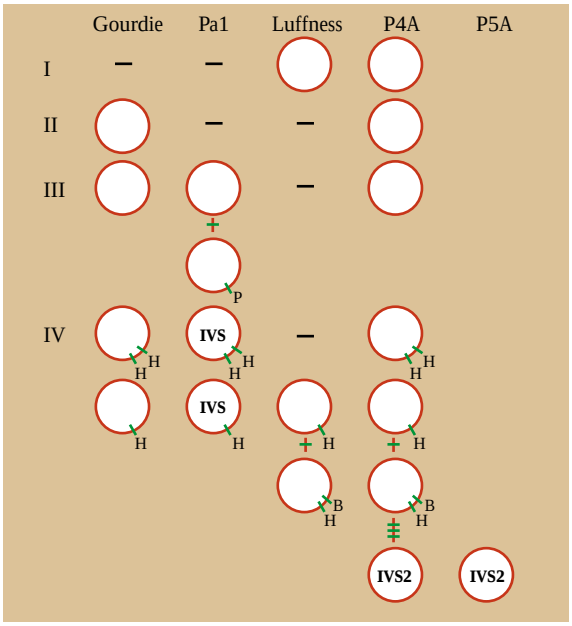

Figure 5 Schematic representation of sequence variants of scmtDNAs I-IV found in populations Gourdie (U.K.), Luffness (U.K.), Pa1 (U.K.), P4A (South America) and P5A (South America).

Implications The novel model of mtDNA structure presented above raises a number of fundamental questions: How is scmtDNA replication in *G. pallida* co-ordinated and gene expression regulated? The respiratory complexes that make up the electron transport chain are encoded as subunits present on both the nuclear and mitochondrial genomes. Increases in mitochondrial biogenesis therefore require co-ordinated increases in gene expression in both the nuclear and mitochondrial compartments. A multipartite mtDNA would impose an additional layer of complexity on this system, as the expression of individual parts of the mtDNA would presumably also have to be co-ordinated. In this regard, it may be significant that each scmtDNA contains an unusually lengthy non-coding region. With the exception of repetitive elements, metazoan mtDNA non-coding sequences are generally far shorter than those found on *G. pallida* scmtDNAs.

How is the essential gene complement of a multipartite genome maintained during both somatic and germ line cell divisions? In contrast with the replication of the nuclear genome, mtDNA replication is unpredictable. While nuclear DNA replication invariably results in the generation of a precise copy of each chromosome, each of the many mtDNA molecules present in a cell is not faithfully copied. As a result, cells containing mixed populations of mtDNA molecules (heteroplasmic cells) can give rise to cells or gametes that contain only one type of mtDNA molecule. In conventional systems this process can be very rapid, with the parental mtDNA genotype being replaced in one generation. With a multipartite mtDNA, it would be necessary to counteract this effect to ensure the viability of progeny.

Other questions concern the mechanism by which an apparently fragmented mtDNA arose and the implications for mtDNA genome evolution. For example, do the components of the *G. pallida* mtDNA recombine and what are the processes that determine the frequency a particular scmtDNA is found at in a population? It is also unclear how widespread this phenomenon is within the Nematoda.