

## Linkage Group V

It should be noted that supercontig 5 represents most but not all of linkage group V as it lacks most of the nucleolus organiser region (NOR) towards the end of the left arm. There is but a single NOR in the genome (McClintock, 1945). The left end of supercontig 5 contains only a short fragment of the NOR containing a single representative of the repeated rRNA unit (18s[ncu15825], 5.8s[ncu15826] and 28s[ncu15827]) of the estimated ~150 repeats of that unit in the NOR (Cox and Peden, 1979, Free et al, 1979). The rRNA unit occurs in the NOR in tandem repeats, and pre-meiotic exchanges generate variations in the number of repeats from 130-172 around that 150 average value (Butler and Metzenberg, 1979).

The current genome assembly contains additional copies of the rRNA repeat unit in supercontigs 8, 9, 12, 15, 16, 18 and 19, hence these also belong to linkage group V. Including these supercontigs together with supercontig 5, the total currently sequenced genome contains ~50 copies of the basic unit, hence approximately one third of the estimated total number. One of these, supercontig 19, also contains subtelomeric sequence at its left end, thus containing the left end of the NOR. Likewise the left end of supercontig 5 contains its right end of the NOR. It is not possible to order supercontigs 8, 12, 15, 16, 18 and 19 within the NOR, but as the NOR contains tandem repeats of the basic rRNA unit and the orientation of that unit is known from the copies in supercontigs 5 and 9, it is possible to identify that supercontigs 8, 15, 16 and 19 are correctly oriented but supercontigs 12 and 18 are reversed, being incorrectly numbered from right to left.

The under-representation of the NOR rRNA sequences in the current assembly, only ~50 copies of the previously estimated ~150 copies, is significant. The published value for the size of the repeating unit is ~9.3kb (Cox and Peden, 1979, Free et al, 1979). However, the actual sequence data now show the unit length to be ~8.8kb. The 50 copies in the current assembly contain ~440kb of NOR, so another ~100 copies or ~880kb is missing. The current stated estimate of genome size of 41.04Mb (actually 41,037,538) which is based only on the total sequenced genome is ~880kb below the best estimate of its true size, ~41.92Mb.

McClintock, B (1945) *Neurospora*. I. Preliminary observations of the chromosomes of *Neurospora crassa*. Am. J. Bot. 32: 671-678.

Cox R A and Peden K (1979), A study of the organisation of the ribosomal ribonucleic acid gene cluster of *Neurospora crassa* by means of restriction endonuclease analysis and cloning in bacteriophage lambda. Mol. Gen. Genet. 174: 17-24.

Free S J, Rice P W and Metzenberg R L (1979) Arrangement of the genes coding for ribosomal ribonucleic acids in *Neurospora crassa*. J. Bacteriol. 137: 1219-1226.

Butler D K and Metzenberg R L (1979) Premeiotic change of nucleolus organiser size in *Neurospora*. Genetics 122: 783-791.