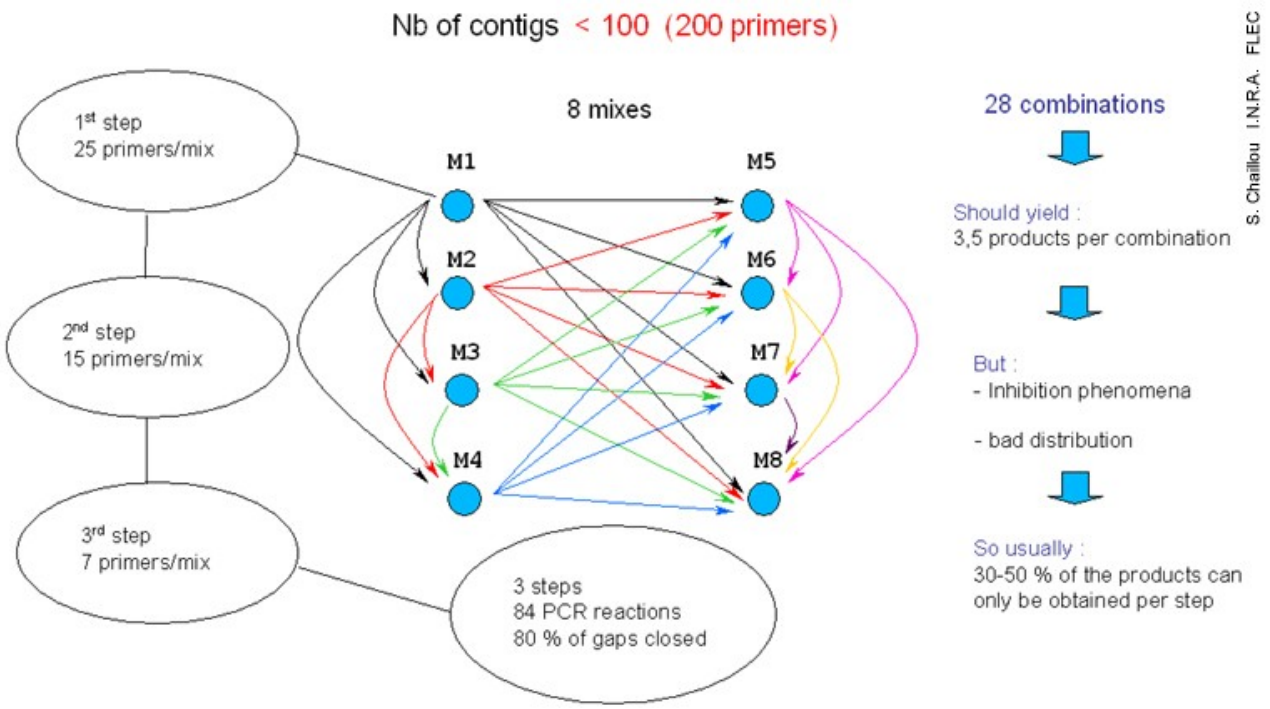
PROTOCOLE USED FOR MULTIPLEX PCR AND SEQUENCING

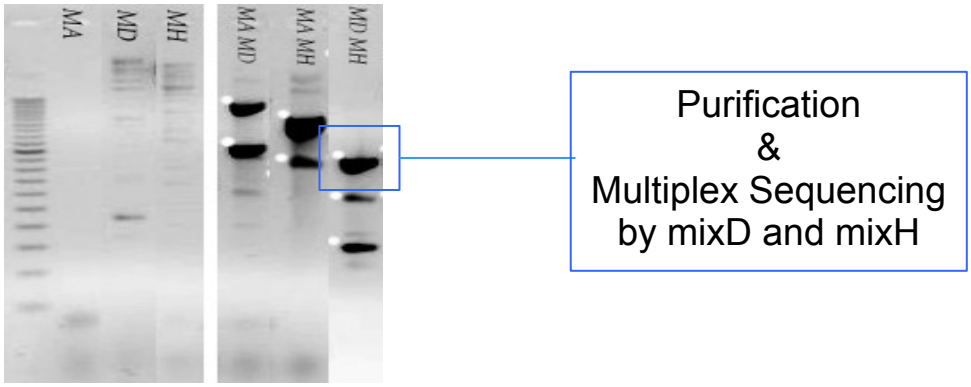
Stéphane Chaillou, April 3rd 2002.

PRINCIPLE

Our Multiplex Long-Range (MLR) PCR is based on the following principle :



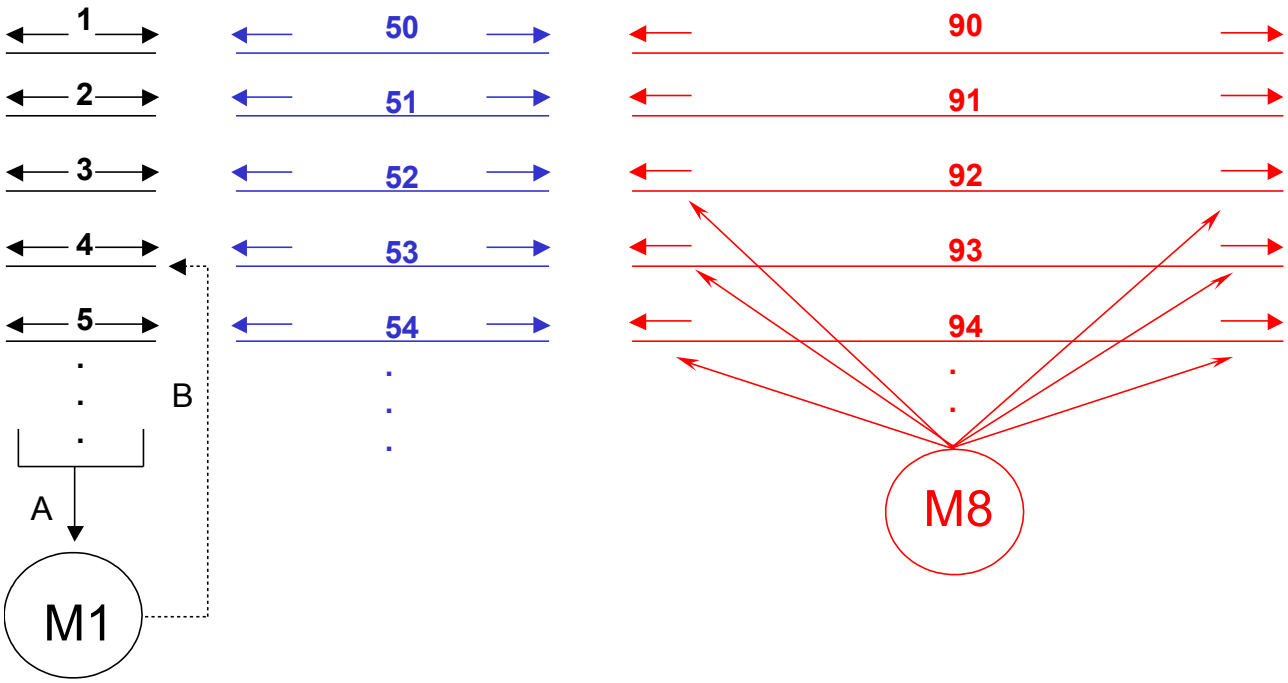
The purpose is to obtain as few products as possible in each M1 ... M8 mixture alone, and as many products as possible in the mixtures combinations (indicated in coloured arrows). This strategy allows to select and purify PCR products which appears only in the combinations. Each product is subsequently sequenced by the corresponding primers mixtures (in which only one primer should be specific for sequencing). Adding of the reads to the shogun assembly will reveal the name of the primers that were responsible for the PCR amplification and will therefore be remove from the batch at the next MLR-PCR step. An example is shown below:



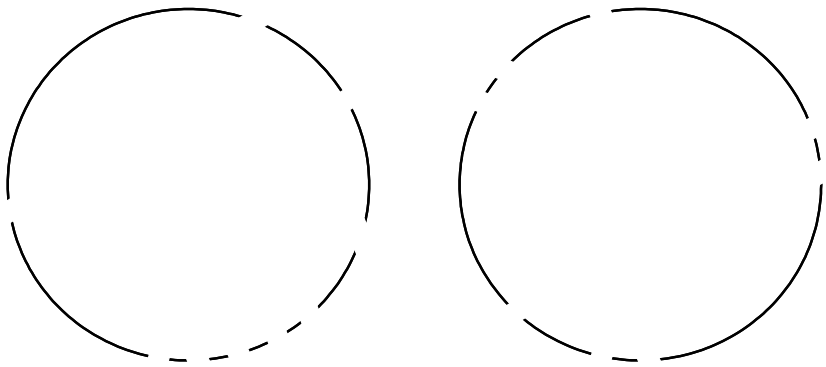
Here, several products were obtained (white dots) by mixture combinations of 3 batches.

CONSTRUCTING THE PRIMERS MIXES

This is of course the critical step. One may thing that with 100 gaps (thus 200 primers), since you don't know the contig's ordering in the final chromosomal map, you just have to randomly choose 25 primers (step 1) out of your 200 ones and hope that it will not give too many products when mixed together. There is a better strategy to improve your choice of primers. With shotgun assemblies, you usually get a mixture of small (1-2 kb), average (10-20 kb) and big contigs (few hundreds kb) with a gradual distribution between the three classes :



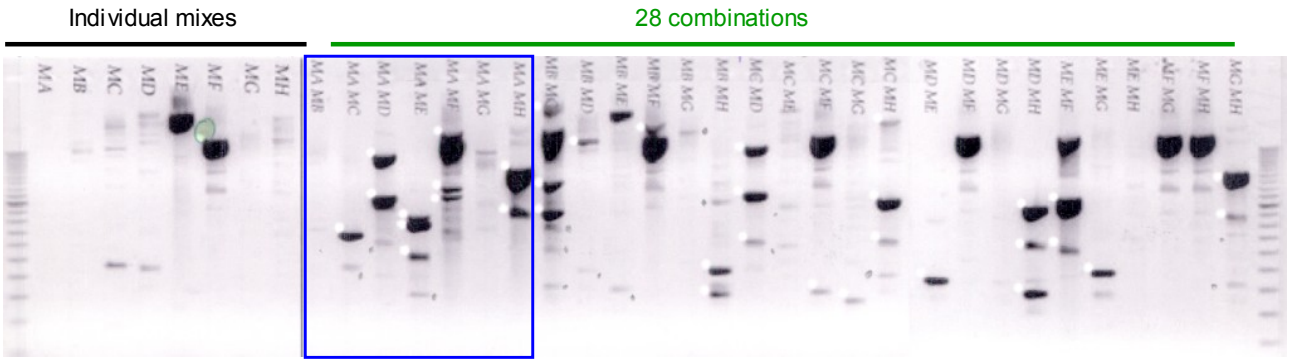
We made the hypothesis that the three classes of contigs may be **randomly distributed** along the chromosome **but not randomly ordered** between them. We estimated that small and average contigs would be mostly located between big contigs which are making a large scaffold :



1 – Possible random distribution 2 – Ordered distribution

This hypothesis revealed to be correct when the *Lactobacillus sakei* 23K chromosome was fully sequenced and the assembly data could be analysed.
Therefore, when creating the mixes, first starts with mixing primers from the same contig (those at both ends) since these two cannot (in principle !) give a PCR product (action A in figure above). Then mix a contig primers pair with other ones from the same class of contigs

(action B in figure above). In such procedure, mix M1 should contain only primers from small contig and mix M8 primers from big contigs. This strategy allows a reduction of PCR amplification of products in mixes and enhance the efficiency of the multiplex when combinations of mixes are applied. In such a strategy, conbination of M1 and M2 should give small amount of products (mixing of small contigs together), such as M7 and M8 (mixing of big contigs together), but any combination in between should give high number of positive products. Below is an example of the second MLR-PCR step of the *L. sakei* 23K gap closure.



Individual mixes give rather few products, only ME and MF show significant amplification of two products (one in each mix). When MA is mixed in its 7 combinations (blue rectangle), it gives no product with MB (also small contigs), one product with MC and more products with mixes of primers from larger contigs. Combinations of primers mixes from average-size contigs (in the middle of the gel) also give rather few products but more are obtained when mixed with MA, MB or MG, MH (mixes from small or big contigs). In this example (step 2) 35 specific products were successfully amplified and sequenced (white dots). The first step has allowed the amplification of 53 products. The remaining 12 products left out of 100 contigs were easily amplified in the third MLR-PCR cycle.