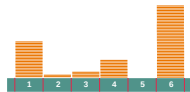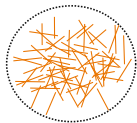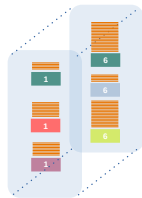Samples are mapped to reference genome and demultiplexed across amplicons

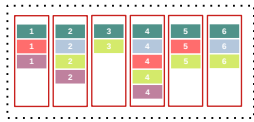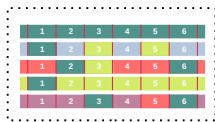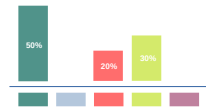Global alignment of reads to references for each amplicon

$$\mathcal{L}(\mathcal{R}|\delta,\pi) = \prod_a^A \prod_r^{R^a} \left(\sum_{g=1}^G P(\mathcal{R}_r^a | \mathcal{S}_g^a, \delta_a) \pi_g \right)$$

E) assign reads to references, knowing proportions and errors
M) estimate proportions and errors

Proportions of references

50%

20% 30%

Reference lineages are cut along amplicon definition and dereplicated