

# Uses of restriction digests

# I: Restriction mapping

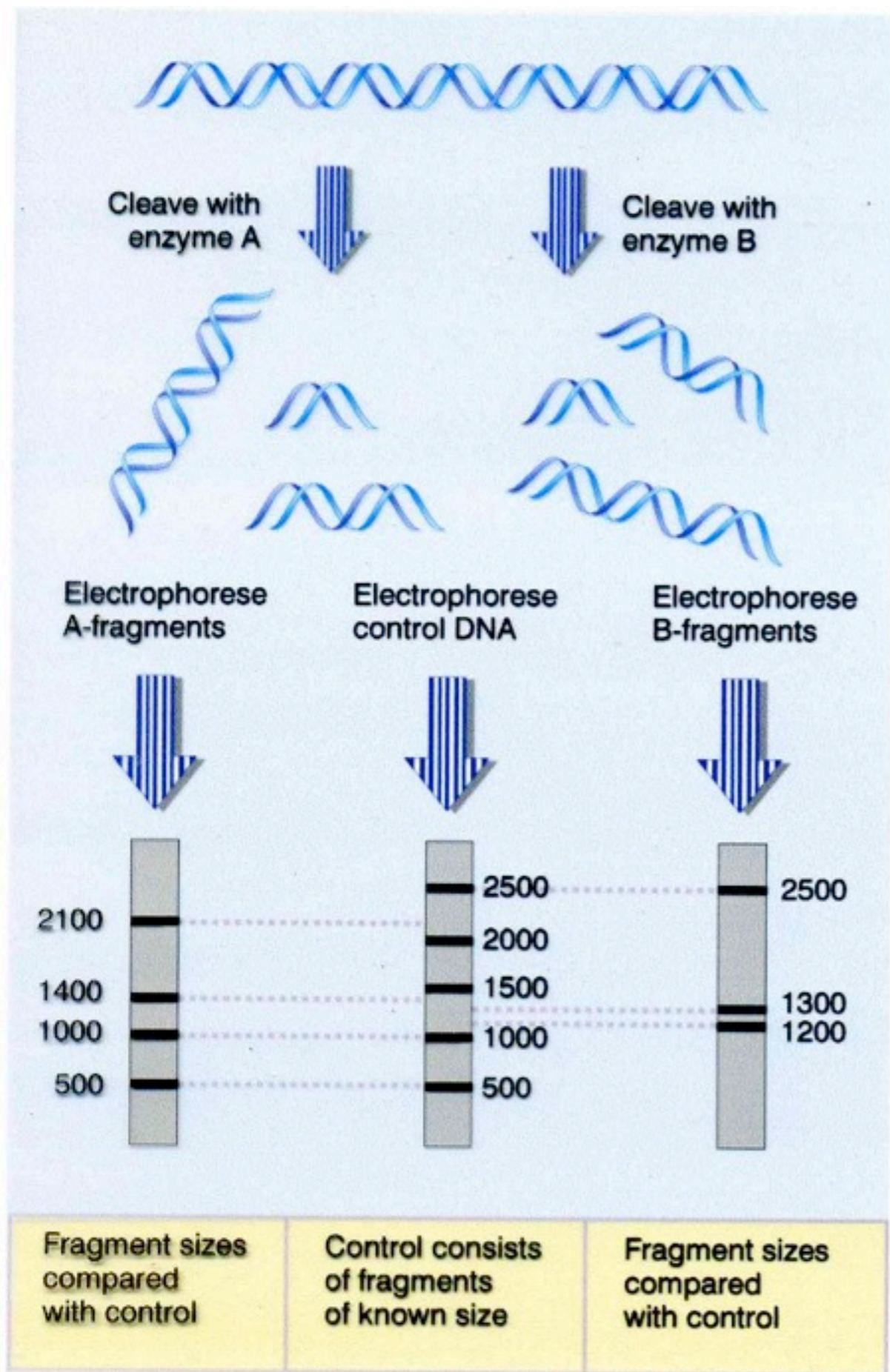
- Because each piece of DNA has a specific set of restriction sites along its length, it will be consistently cut into pieces of repeatable length by each enzyme. Comparing the pieces left by each enzyme and pair of enzymes usually allows you to produce a unique 'map' of the DNA molecule.

# Variation in fragments

- The number and size of fragments depends on
  - whether the DNA is linear or circular
  - the number of cut sites
  - the relative position of the sites to each other

# Potential problems

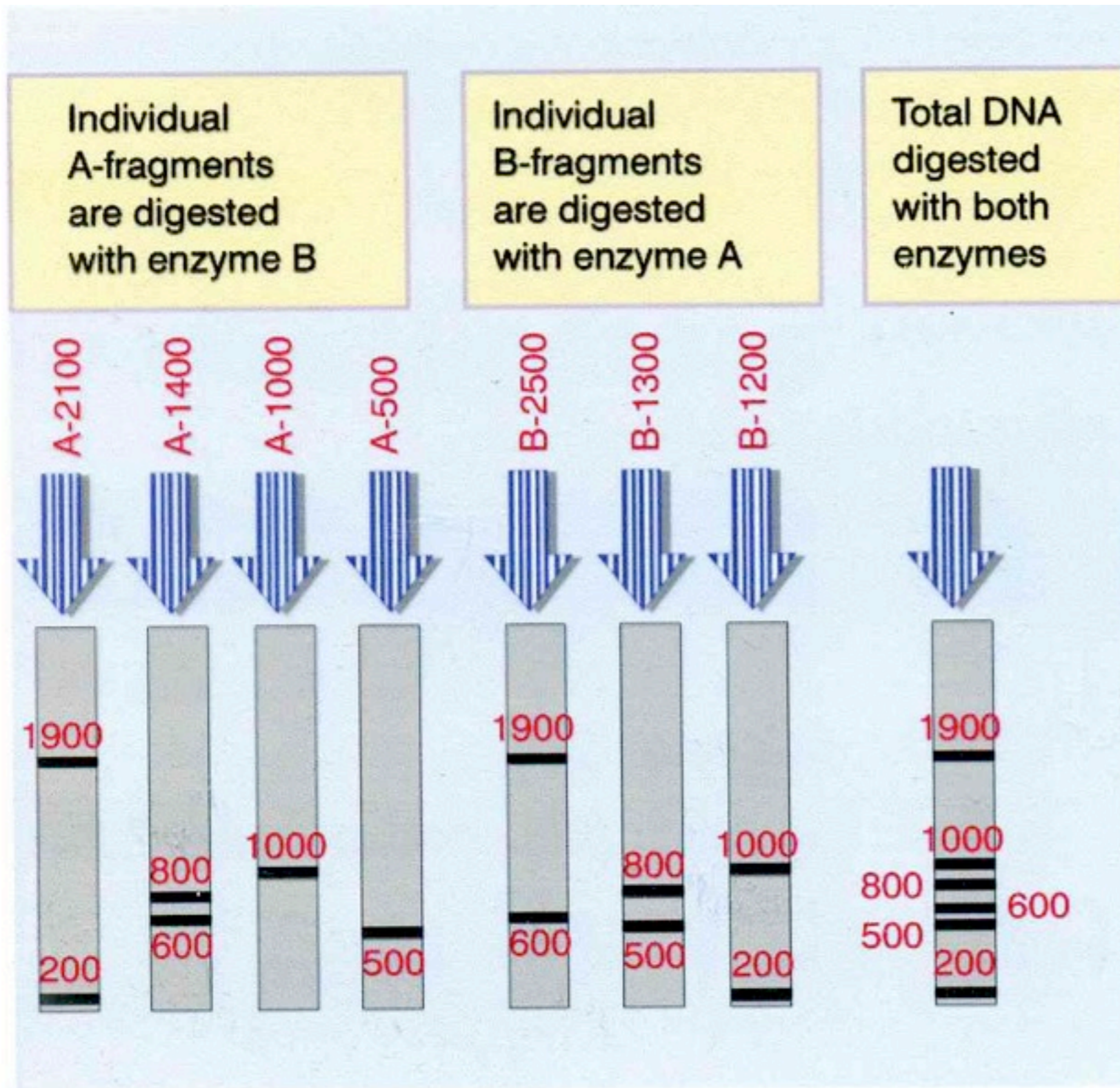
- If two cut sites are very close the fragment may be so small that it is invisible, or runs off the gel
- If two fragments are very similar in size they may appear as a single band
- However: in most cases you get an unambiguous pattern of bands



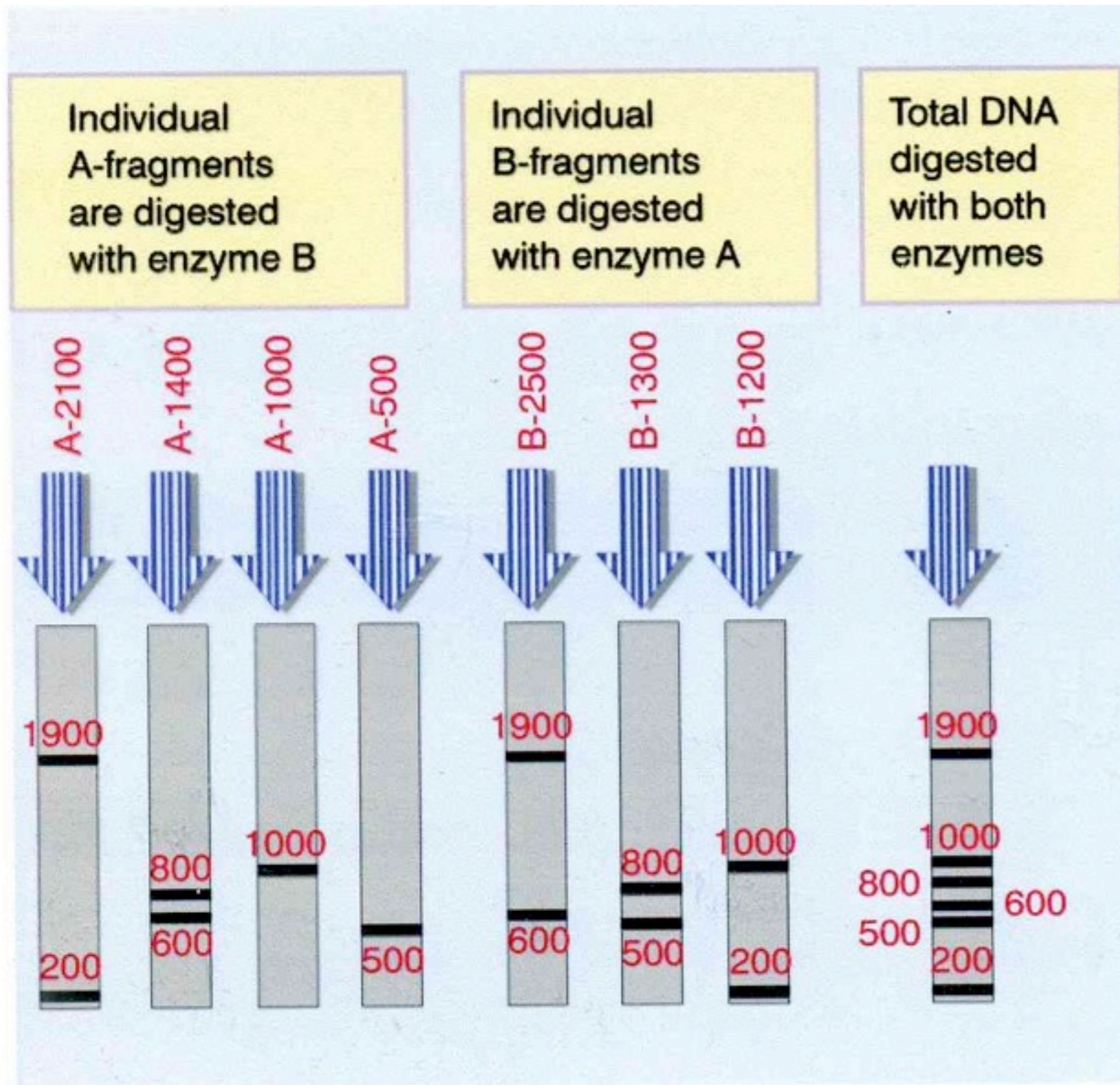
Here a piece of DNA has been completely digested by two enzymes and each digest run on a gel. In each case the total length of the fragments is 5000 bp but at this stage you cannot deduce anything about the arrangement of the individual cut sites

# Double digests

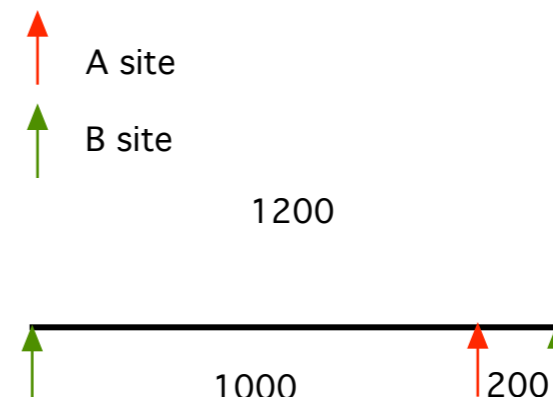
- By digesting the DNA with both enzymes, and comparing the results to the single digests, it is usually possible to deduce the relative location of each of the restriction sites.
- Here each of the fragments from the A digest are purified and recut with B and vv. In addition a combined digest is done and run for comparison.



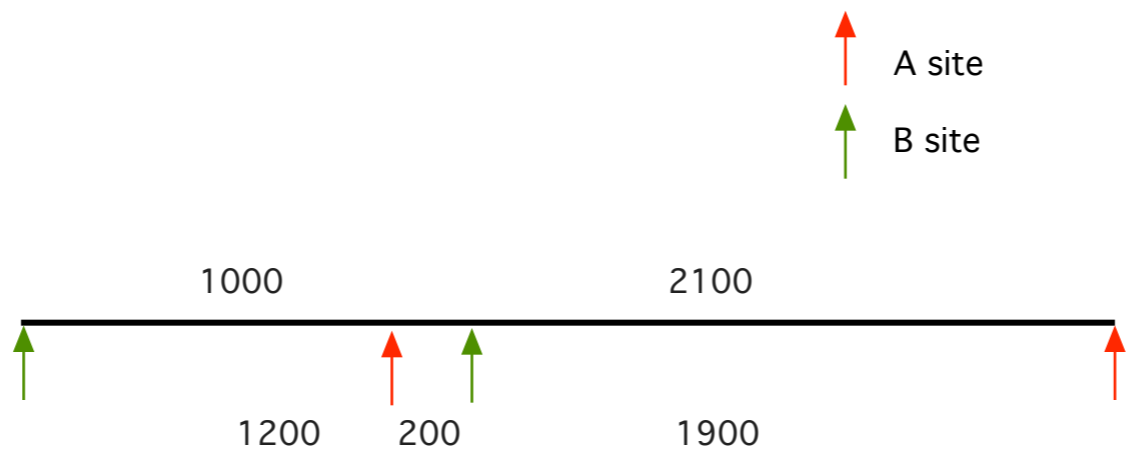
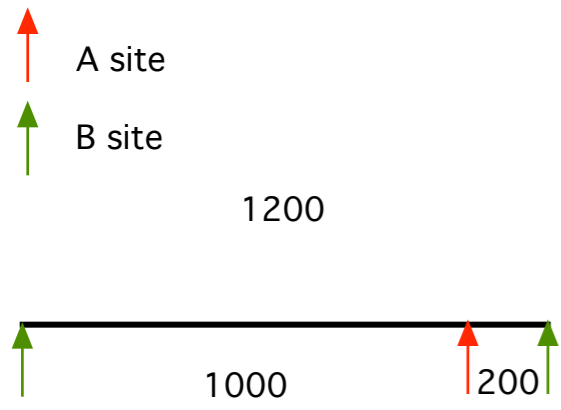
Looking at each of these results it is possible to work out their relationship to each other. Note every band in the redigested A set is also present in the B set: in both cases they **must** end up producing the same result



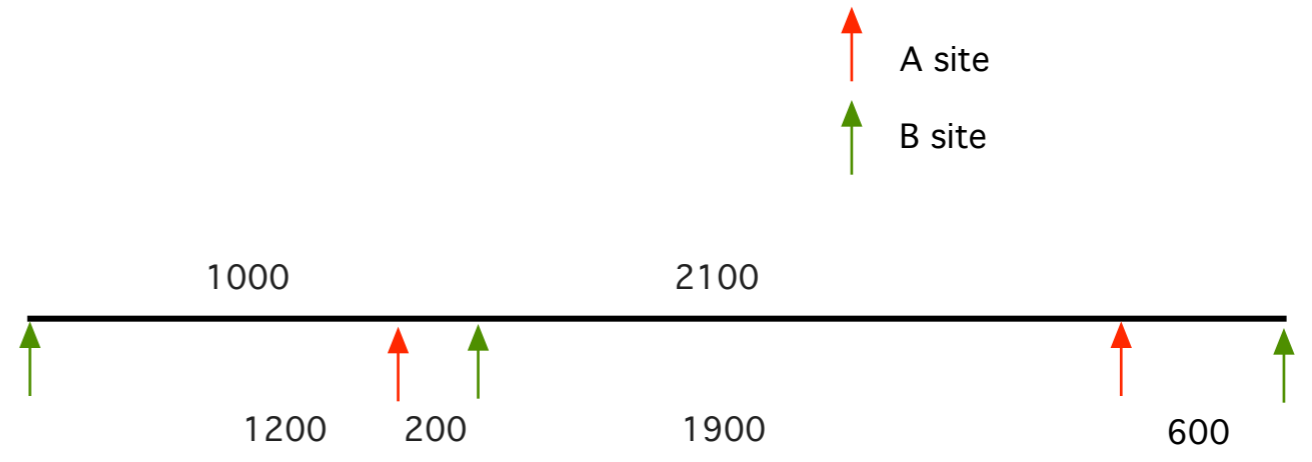
Where do you start? It really doesn't matter: suppose we take the last one: a 1200 bp fragment from the B digest get cut into a 200 and a 1000 bp fragment by A. So, it must look like this:



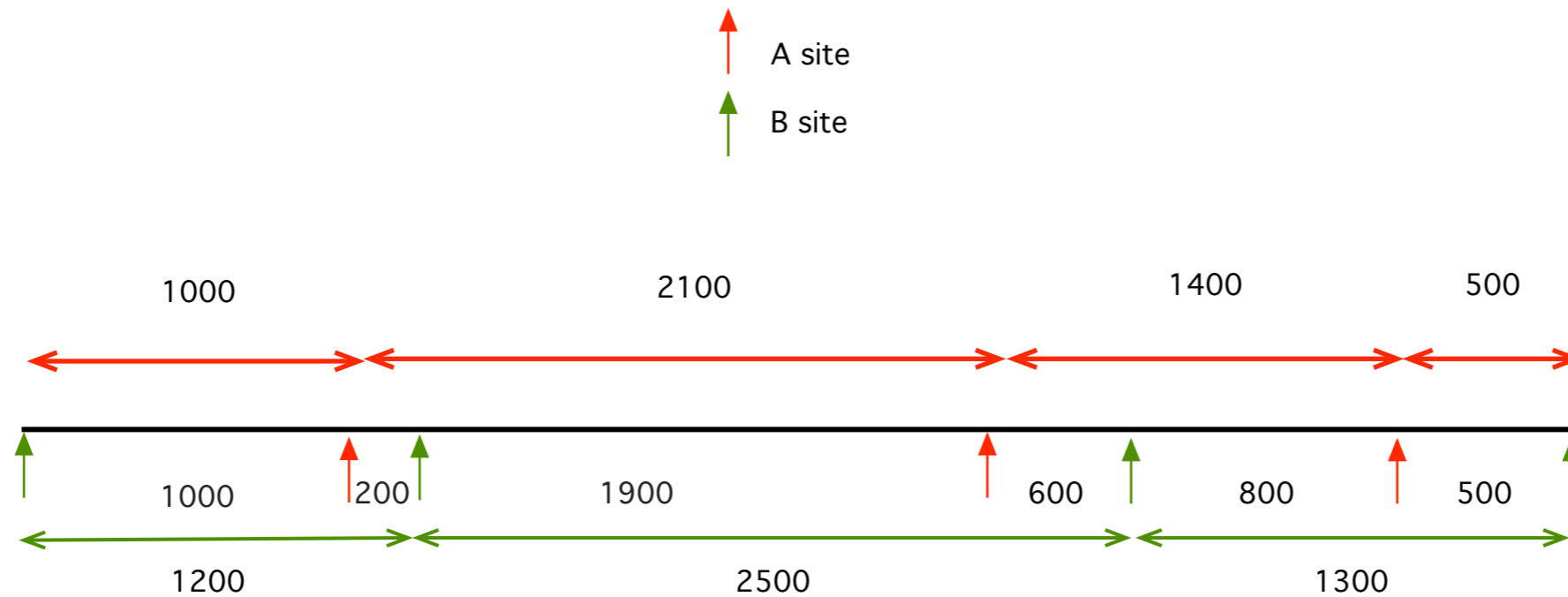
The 200 bp fragment must be the same one that we see in the first A double digest: so it must be attached to the 1900 bp fragment



and this, in turn, must be attached to the 600 bp fragment



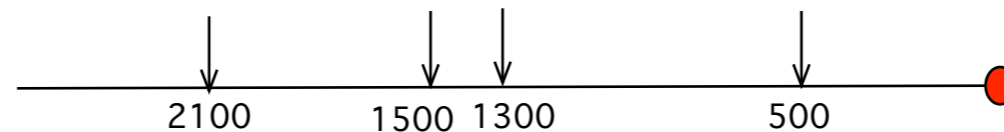
Completing the map gives the following. This shows how the interaction between the two enzymes gives the observed fragments



In this case the DNA is linear, so the ends can be thought of as being both a A and B cut site. If it were circular, with the two ends coming from a B site, then the 1000 and 500 A pieces would be a single 1500 bp piece

# Mapping by end-labeling

- An alternate mapping method is to label one end of the DNA and carry out partial digests



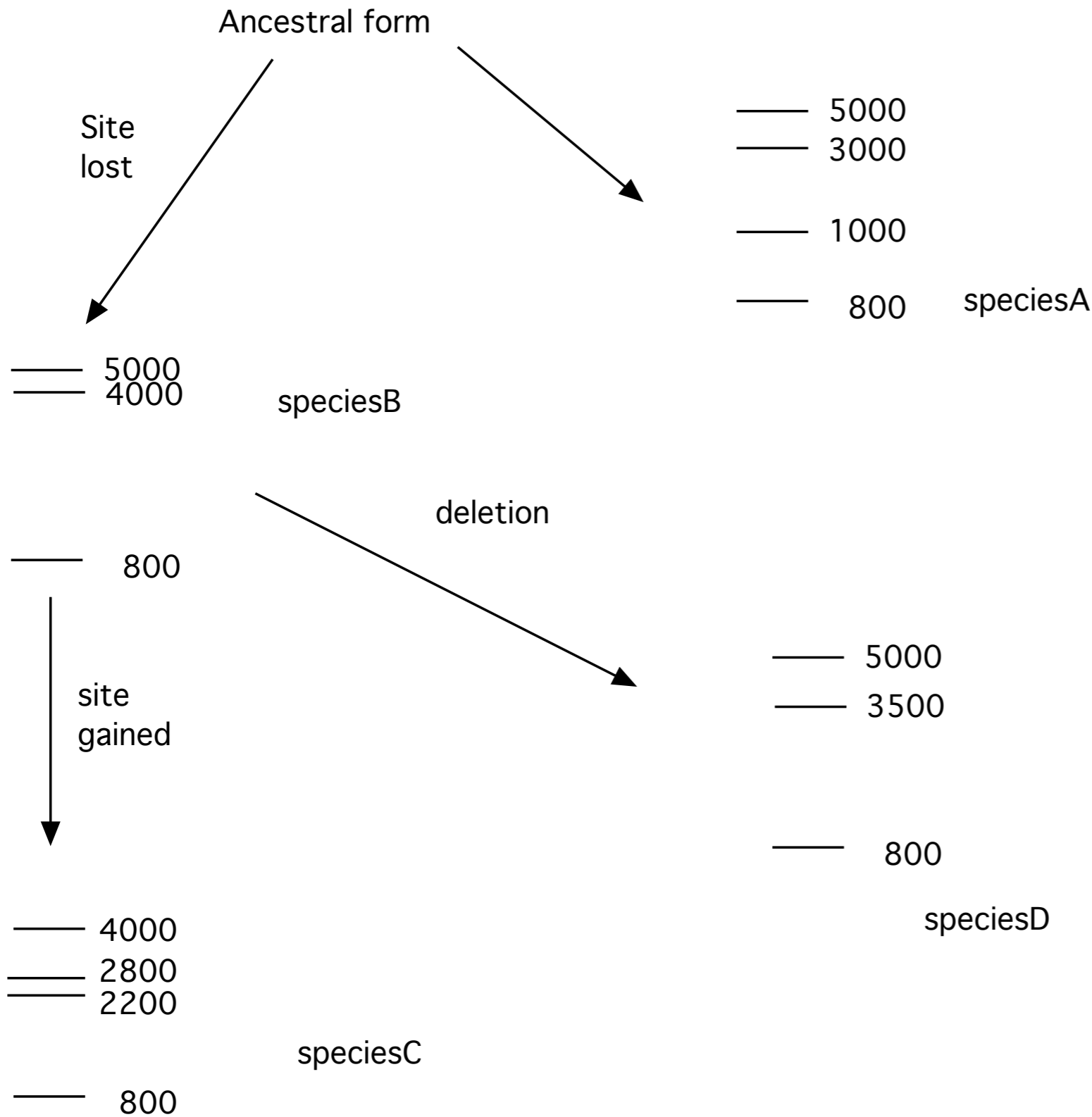
The partial digest produces all possible size fragments but the only ones that are visible are those with a labeled end. These will be 500, 1300, 1500, 2100 and 2500 bp long, indicating the location of each of the restriction sites.

# RFLPs

- If DNA is extracted from closely related organisms, and treated with the same restriction enzyme(s) the resultant restriction fragments should be identical, or at least very similar.
- A number of small (and not so small) variations in the number and size of bands. These variations are termed Restriction Fragment Length Polymorphisms (RFLPs)

# Origin of RFLPs

- If two DNA molecules start out the same, they can produce different patterns because of:
  - Gain or loss of a restriction site. This can be due to an alteration of a single base. The total length of the molecule stays the same, but the number of fragments changes
  - A significant addition or deletion of DNA between sites. Unless a site is duplicated or deleted the number of fragments stays the same, but one of the bands changes in size.



You usually look at all of the species and try to decide which must have given rise to the other by the simplest route. For example: A cannot have led directly to C by a single step. In this case we can't tell if A came from B or an ancestral form. Looking at RFLPs using another enzyme may help.

# Limitations

- Although RFLP analysis has been very valuable, it is limited to relatively small genomes: plasmids, viruses and organelles. Once the number of fragments gets above 15-20 it becomes very difficult to analyze the results, even with computer programs to sort through all possible solutions.
- With eukaryotic genomes the number of fragments reach into the thousands and overlap to the point where no individual bands are visible. In this case a different approach is needed.

# DNA Fingerprinting

- This technique allows the DNA from two samples to be analyzed for similarity.
- It relies on the fact that the likelihood of two independent events is the product of their individual probabilities.
- For example: the odds of rolling two sixes is  $1/36$ , the odds of picking the ace of spades is  $1/52$ . The odds of doing both is  $1/1872$

- Both samples are initially cut with the same restriction enzyme. If this is an 8 cutter it will (on average) cut every 65,000 bp. In a sample of human DNA that means about 46,000 fragments. Since the sites are randomly arranged, the actual sizes of the fragments will range from a few bp to over a million bp. Even so, there are so many fragments they will form a continuous overlapping smear down the gel.

- If you **were** able to see and identify every band then almost certainly everyone would have a unique pattern. Just as with RFLPs the odds of every site being in exactly the same position in two samples is virtually zero.
- However: we can't see all the bands, so another method of analysis is needed. The solution is to look for the pattern of repeated sequences present in the sample.

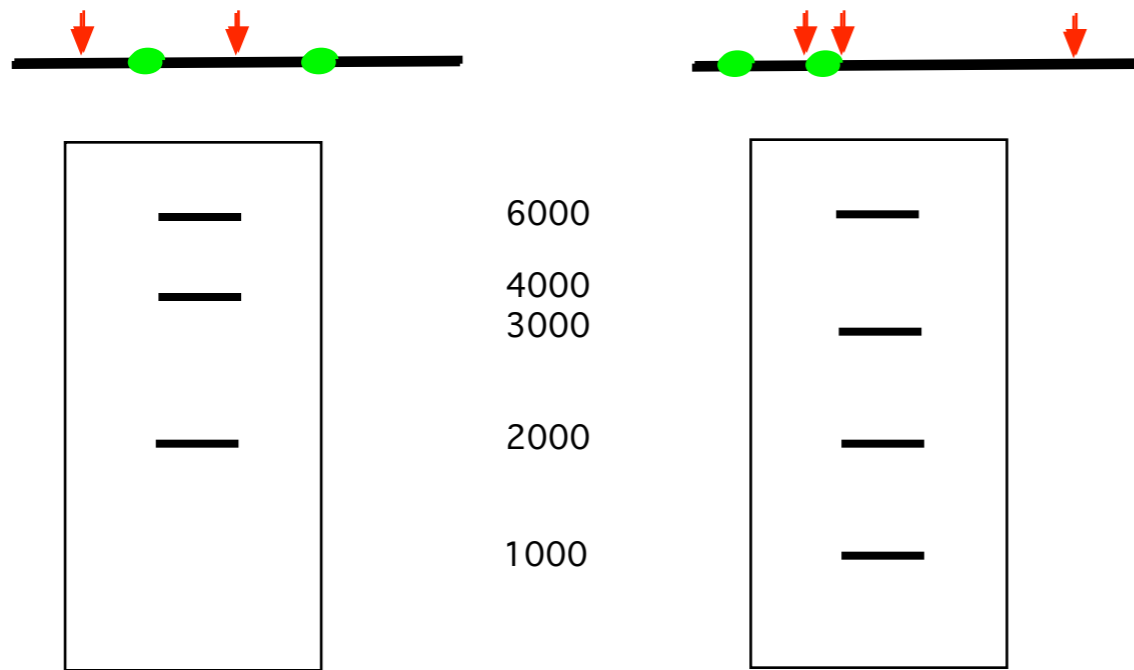
- Some repeated sequences occur just 10-20 times in the genome. The number of repeats is random, as is their location in the genome. This means that each person will have a specific *number* of their restriction fragments containing a copy of the repeated sequence.
- *Which* of the restriction fragments carry the repeats depends on *where* the repeats were in the genome and *where* the restriction sites were. These are all independent of each other

# Identifying the location of repeats

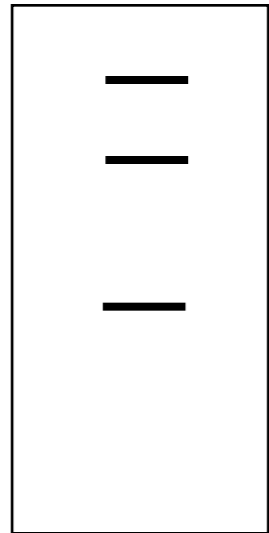
- After the DNA has been digested and run on a gel it is blotted to a filter under denaturing conditions, so it is in single-stranded form.
- The filter is incubated with labeled copies of a probe that will only bind to the repeated sequence. After washing, the location of the repeated sequences between samples is compared



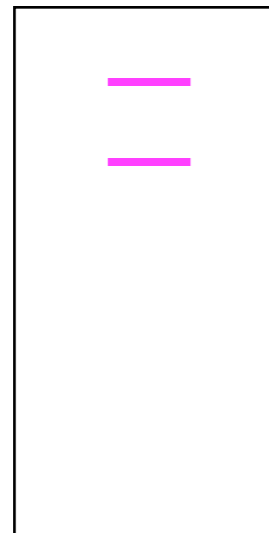
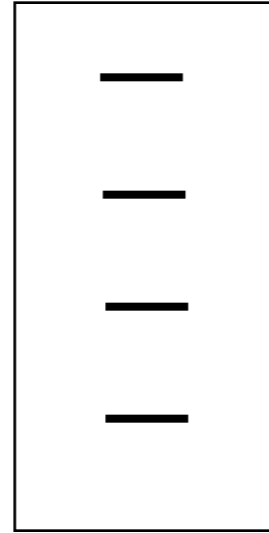
Two samples of DNA, one with two restriction sites, the other with three. Each has two repeated sequences at different locations



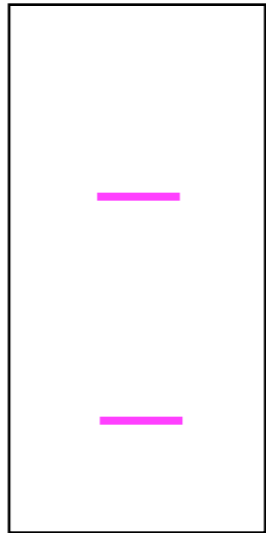
When cut and run on a gel, each sample produces a set of bands. Remember: in reality there would be thousands of overlapping bands.



6000  
4000  
3000  
2000  
1000



6000  
4000  
3000  
2000  
1000



Each is treated with a fluorescent probe and examined under UV. Only those bands that contain the probe light up.

If the patterns match the process can be repeated with different enzymes and probes to confirm the process.