

UNIX pipes and redirection

Redirection within UNIX allows a user to control where input comes from and output goes to. Pipes allow a user to pass the output from one command directly to the input of another command.

Redirection

>

The > symbol redirects the output from standard out (This is what you see on the screen) to a file. **Beware** if the file already exist, it will be erased and overwritten without warning, so be careful.

An example of this is for the commonly used mapping tool bwa. By default `bwa mem` will print the output to the screen. However you can write this to a file using >

```
bwa mem reference.fasta sample1_R1.fastq.gz sample1_R2.fastq.gz > sample1.sam
```

This command writes the output from `bwa mem` (in SAM format) to a file called `sample1.sam`

>>

The >> symbol redirects the output from standard out to a file but rather than overwriting the file, appends the output to the existing file. If the file does not exist it will just create a new one

<

The < symbol takes input from a file. An example below using the sickle trimming software that does not take gzipped fastqs is as follows

```
sickle se -f <(gunzip -c sample1_R1.fastq.gz) -t sanger -o trimmed.fa
```

Here `gunzip -c` writes the unzipped output to the STDOUT and rather than write to disc we convert it on the fly and read it into sickle

Piping |

The | symbol passed the out from one command onto another for further processing. An example using `bwa mem` is

```
bwa mem reference.fasta sample1_R1.fastq.gz sample1_R2.fastq.gz | samtools view -bS -F 4 | samtools sort -o sample1.sorted.bam
```

Here rather than running `bwa mem`, followed by `samtools view` to convert to a bam file keeping only mapped reads (-F 4) and then sorting, these can be joined together through the pipe command into a single line. This has the advantage that the intermediate file are not written to disc.