

```

## Regular expressions
## reference: http://perldoc.perl.org/perlref.html

open(IN,"<file.fasta") or die "ERROR! I cannot read the file\n";
$num_seq=0;
while($line=<IN>){
    if($line=~/^>/){
        $num_seq++;
    }
}
close(IN);
print "There are $num_seq sequences\n";

## pattern matching
@oligos=("atgc","atgcaa","atggaat");
foreach $oligo (@oligos){
    if($oligo=~^/atgc/){
        print $oligo."\n";
    }
}

@oligos=("atgc","atgcaa","atggaat");
foreach $oligo (@oligos){
    if($oligo=~^/atg[cg]/){
        print $oligo."\n";
    }
}

@oligos=("atgc","atgcaa","atggaat");
foreach $oligo (@oligos){
    if($oligo=~^/atg[^c]/){
        print $oligo."\n";
    }
}

@oligos=("atgc","atgcaa","atggaat");
foreach $oligo (@oligos){
    if($oligo=~^/gc$/){
        print $oligo."\n";
    }
}

# quantifiers
$seq="cgatcagctacgatcgatcagcatcgatcaaaaaaaaaaaaaaaaaaaaaa";
if(uc($seq)=~/A{3}$/){
    print "PolyA\n";
}
else{
    print "No polyA\n";
}

# capturing groups
$line="__CGATCGATCGAT__";
$line=~/_([ATGC]+)_/;
print $1."\n";

#substitution
$seq="CGATCGAGCTACGATCAGCTACGATC";
$seq=~s/AT/**/;
print $seq;

$seq="CGATCGAGCTACGATCAGCTACGATC";
$seq=~s/AT/**/g;
print $seq;

@files=`ls|grep .fasta`;
chomp(@files);

```

```
foreach $file (@files){
  open(IN,"<$file");
  $description=<IN>;
  print "File: $file; Decription: ";
  if($description=~/^>([A-Za-z1-9]+)[ \t\n]/){
    print $1."\n";
  }
  else{
    print "no description\n";
  }
  close(IN);
}
```