class 6

Anqi Feng

my first function

```
add<-function(x,y=1){
  x+y
}
```

can i just use this?

```
add(10,10)
```

```
[1] 20
```

```
add(10)
```

```
[1] 11
```

```
add(1,1)
```

```
[1] 2
```

```
add(x=1,y=100)
```

```
[1] 101
```

```
add(c(100,1,100),1)
```

```
[1] 101   2 101
```

```
add(10)
```

```
[1] 11
```

Q;make a function: "generate_DNA()" that makes a random nucleotide sequence of any length

```
bases <- c("A", "C", "G", "T")
sequence <- sample(bases, size=5, replace=TRUE)
```

That is my working lines, now i can make it into a function

```
generate_dna<-function(length){
  bases <- c("A", "C", "G", "T")
sequence <- sample(bases, size=length, replace=TRUE)
return(sequence)
}
```

```
generate_dna(10)
```

```
 [1] "C" "G" "T" "G" "G" "G" "T" "G" "G" "T"
```

```
bio3d::aa.table
```

```
    aa3 aa1    mass       formula                      name
ALA ALA   A  71.078    C3 H5 N O1                   Alanine
ARG ARG   R 157.194  C6 H13 N4 O1                  Arginine
ASN ASN   N 114.103   C4 H6 N2 O2                 Asparagine
ASP ASP   D 114.079    C4 H4 N O3              Aspartic Acid
CYS CYS   C 103.143  C3 H5 N O1 S                    Cystein
GLN GLN   Q 117.126   C4 H9 N2 O2                 Glutamine
GLU GLU   E 128.106    C5 H6 N O3              Glutamic Acid
GLY GLY   G  57.051    C2 H3 N O1                    Glycine
HIS HIS   H 137.139   C6 H7 N3 O1                  Histidine
ILE ILE   I 113.158   C6 H11 N O1                 Isoleucine
LEU LEU   L 113.158   C6 H11 N O1                    Leucine
LYS LYS   K 129.180  C6 H13 N2 O1                     Lysine
MET MET   M 131.196  C5 H9 N O1 S                 Methionine
PHE PHE   F 147.174    C9 H9 N O1              Phenylalanine
```

```
PRO PRO   P  97.115    C5 H7 N O1                         Proline
SER SER   S  87.077    C3 H5 N O2                         Serine
THR THR   T 101.104    C4 H7 N O2                       Threonine
TRP TRP   W 186.210 C11 H10 N2 O1                       Tryptophan
TYR TYR   Y 163.173    C9 H9 N O2                         Tyrosine
VAL VAL   V  99.131    C5 H9 N O1                         Valine
ABA ABA   X  85.104   C4 H7 N1 O1          alpha-aminobutyric acid
ASH ASH   D 115.087    C4 H5 N O3            Aspartic acid Neutral
CIR CIR   R 157.170  C6 H11 N3 O2                        citrulline
CME CME   C 179.260 C5 H9 N O2 S2 s,s-(2-hydroxyethyl)thiocysteine
CMT CMT   C 115.154  C4 H5 N O1 S                  o-methylcysteine
CSD CSD   C 134.134  C3 H4 N O3 S          s-cysteinesulfinic acid
CSO CSO   C 119.142  C3 H5 N O2 S                s-hydroxycysteine
CSW CSW   C 135.142  C3 H5 N O3 S               cysteine-s-dioxide
CSX CSX   C 119.142  C3 H5 N O2 S                   s-oxy cysteine
CYM CYM   C 102.135  C3 H4 N O1 S                 Cystein Negative
CYX CYX   C 102.135  C3 H4 N O1 S                   Cystein SSbond
DDE DDE   H 280.346 C13 H22 N5 O2                      diphthamide
GLH GLH   E 129.114    C5 H7 N O3           Glutatmic acid Neutral
HID HID   H 137.139   C6 H7 N3 O1                        Histidine
HIE HIE   H 137.139   C6 H7 N3 O1                        Histidine
HIP HIP   H 138.147   C6 H8 N3 O1               Histidine Positive
HSD HSD   H 137.139   C6 H7 N3 O1                        Histidine
HSE HSE   H 137.139   C6 H7 N3 O1                        Histidine
HSP HSP   H 138.147   C6 H8 N3 O1               Histidine Positive
IAS IAS   D 115.087    C4 H5 N O3                     beta-aspartyl
KCX KCX   K 172.182  C7 H12 N2 O3        lysine nz-carboxylic acid
LYN LYN   K 129.180  C6 H13 N2 O1                   Lysine Neutral
MHO MHO   M 147.195  C5 H9 N O2 S                 s-oxymethionine
MLY MLY   K 156.225  C8 H16 N2 O1                n-dimethyl-lysine
MSE MSE   M 178.091 C5 H9 N O1 SE                selenomethionine
OCS OCS   C 151.141  C3 H5 N O4 S            cysteinesulfonic acid
PFF PFF   F 165.164  C9 H8 F N O1         4-fluoro-l-phenylalanine
PTR PTR   Y 243.153 C9 H10 N O5 P                o-phosphotyrosine
SEP SEP   S 167.057  C3 H6 N O5 P                    phosphoserine
TPO TPO   T 181.084  C4 H8 N O5 P                 phosphothreonine
```

```
unique(bio3d::aa.table$aa1)[1:20]
```

```
 [1] "A" "R" "N" "D" "C" "Q" "E" "G" "H" "I" "L" "K" "M" "F" "P" "S" "T" "W" "Y"
[20] "V"
```

```r
generate_protein<-function(l){
  aas<-unique(bio3d::aa.table$aa1)[1:20]
  poly <- sample(aas, size=l, replace=TRUE)
  poly<-paste(poly, collapse = "")
  return(poly)
}
```

```r
generate_protein(10)
```

```
[1] "CGDQVMISGA"
```

Generate protein sequences of length 6 to 12

```r
answer<-sapply(6:12, generate_protein)
```

```r
cat(paste( ">id.", 6:12, "\n", answer,sep="", "\n"))
```

```
>id.6
PRGSSC
 >id.7
PYFDTES
 >id.8
FIAYDKKH
 >id.9
EARHTDTRS
 >id.10
PMKHDTISIE
 >id.11
LAYDKVTPQGK
 >id.12
RVNCEECWAEAG
```