

Løsningsforslag

```
def les_inn_dna_sekvenser():
```

```
    friske = []
```

```
    syke = []
```

```
    for i in range(0, 20):
```

```
        nummer = str(i)
```

```
        fil = open("healthy_" + nummer + ".txt")
```

```
        sekvens = fil.read()
```

```
        friske.append(sekvens)
```

```
        fil.close()
```

```
        fil = open("diseased_" + nummer + ".txt")
```

```
        sekvens = fil.read()
```

```
        syke.append(sekvens)
```

```
        fil.close()
```

```
    return friske, syke
```

```
def finn_fordeling(sekvenser):
```

```
    fordeling = []
```

```
    # Lag en tom liste med en teller for A,C,G,T på hver posisjon
```

```
    for i in range(len(sekvenser[0])):
```

```
        fordeling.append([0, 0, 0, 0])
```

```
    for i in range(0, len(sekvenser[0])):
```

```
        for sekvens in sekvenser:
```

```
            if sekvens[i] == "A":
```

```
                fordeling[i][0] += 1
```

```
            elif sekvens[i] == "G":
```

```
                fordeling[i][1] += 1
```

```
            elif sekvens[i] == "C":
```

```
                fordeling[i][2] += 1
```

```
            elif sekvens[i] == "T":
```

```
                fordeling[i][3] += 1
```

```
    return fordeling
```

```
def finn_forskjeller(friske_fordeling, syke_fordeling):
```

```
    for i in range(0, len(friske_fordeling)):
```

```
        if friske_fordeling[i] != syke_fordeling[i]:
```

```
            print("!!! Forskjell på posisjon ", i)
```

```
            print(friske_fordeling[i])
```

```
            print(syke_fordeling[i])
```

```
friske, syke = les_inn_dna_sekvenser()
friske_fordeling = finn_fordeling(friske)
syke_fordeling = finn_fordeling(syke)
finn_forskjeller(friske_fordeling, syke_fordeling)
```

Riktig posisjon skal være 951946, dvs 951947 i genomet.