Code to turn data into text files

```
Wgenesample =
  StringJoin["PCV1num", ToString[InputString["What number individual is this?"]]];
basepairs = ToString[
   {InputString["Paste the base pair sequence (ex: AAGCTATGG ) here"]}];
source = ToString[InputString["What's the source? (ex: GenBank: AB043895.5)"]];
(*SpecialNote=ToString[InputString["Any Special Notes? If not type 'no'. "]];*)

(* OtherInput = ToString[InputString[" Enter Prompt for OtherInput Here "]];*)
```

```
lettersample = {basepairs} // ToString;
LetterDNAtoNum[Sample_] := ToExpression[StringReplace[ToString[
     {StringReplace[StringReplace[ToString[{Sample}], {"," → "", " " → "", "{" → "",
         "}" → "", "(" → "", ")" → "", "[" → "", "]" → "", ";" → "", ":" → "", "_" → "",
         "+" → "", "&" → "", "/" → "", "." → "", "RowBox" → "", "Null" → "", "
        " → "", "
" → ""}], {"0" → "0,", "1" → "1,", "2" → "2,",
        "3" → "3,", "A" → "0,", "C" → "1,", "G" → "2,", "T" → "3,", "a" → "0,",
        "c" → "1,", "g" → "2,", "t" → "3,", "U" → "3", "u" → "3", "N" → ""}]}
    ], ",}" → "}"]]

(* N Removed by/in the above code *)
numgenesample = LetterDNAtoNum[lettersample];
lengthofgeneitself = Length[Flatten[numgenesample]];
M = numgenesample;
```

To produce a .txt file of the gene run the following (grey) cell

To open the created file include *SystemOpen[txtfilename]*

```
(*txtfilename=
   StringReplace[StringJoin[StringReplace[StringJoin[Wgenesample, " ", source],
      {"gene"→ "",","."→ "_","," "→ ""}],".txt"],{"GenBank:"→ "gb"}];
 PCV1directory="C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals\\";
 Export[FileNameJoin[{PCV1directory,txtfilename}],Flatten[numgenesample]]
(*Sends the output to a specific file *)

(*Export[txtfilename,Flatten[numgenesample]]*)
(*Print["This produced a .txt file with the name ", txtfilename]
 SystemOpen[txtfilename]*)*)
```

```
C:\Users\George\Documents\SVD DNA stuff\PCV1 individuals\PCV1num50gbDQ358813_1.txt
```

Reading the data into lists

```
PCV1directory = "C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals\\";
```

```
ReadList["C:\\Users\\George\\Documents\\SVD
    DNA stuff\\PCV1 individuals\\PCV1num1gbKX816645_1.txt"];
```

```
StringJoin[PCV1directory, "PCV1num", "k", "gbKX816645_1.txt"]
```

```
C:\Users\George\Documents\SVD DNA stuff\PCV1 individuals\PCV1numkgbKX816645_1.txt
```

```
ReadList[StringJoin[PCV1directory, "PCV1num", "2", ToString[___ ___ ___ ___], ".txt"]]
```

```
$Failed
```

```
PCV1ListofLists =
 Import["C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals"]
```

```
{PCV1num10gbKC924758_1.txt, PCV1num11gbKC894933_1.txt,
 PCV1num12gbKC878437_1.txt, PCV1num13gbKC733436_1.txt, PCV1num14gbJX566507_1.txt,
 PCV1num15gbKC447455_1.txt, PCV1num16gbAY099501_1.txt, PCV1num17gbJN133303_1.txt,
 PCV1num18gbJN133302_1.txt, PCV1num19gbJN398656_1.txt, PCV1num1gbKX816645_1.txt,
 PCV1num20gbGU799575_1.txt, PCV1num21gbHM143844_1.txt, PCV1num22gbU49186_1.txt,
 PCV1num23gbAY219836_1.txt, PCV1num24gbAY184287_1.txt, PCV1num25gbGU722334_1.txt,
 PCV1num26gbGU371908_1.txt, PCV1num27gbDQ650650_1.txt, PCV1num28gbDQ659154_1.txt,
 PCV1num29gbDQ659153_1.txt, PCV1num2gbKJ808815_1.txt, PCV1num30gbDQ494788_1.txt,
 PCV1num31gbDQ494787_1.txt, PCV1num32gbDQ472016_1.txt, PCV1num33gbDQ472015_1.txt,
 PCV1num34gbDQ472014_1.txt, PCV1num35gbDQ472013_1.txt, PCV1num36gbDQ472012_1.txt,
 PCV1num37gbAY699796_1.txt, PCV1num38gbAY660574_1.txt, PCV1num39gbFJ475129_2.txt,
 PCV1num3gbKJ746930_1.txt, PCV1num40gbY09921_1.txt, PCV1num41gbAF012107_1.txt,
 PCV1num42gbFJ159693_1.txt, PCV1num43gbFJ159692_1.txt, PCV1num44gbFJ159691_1.txt,
 PCV1num45gbFJ159690_1.txt, PCV1num46gbFJ159689_1.txt, PCV1num47gbEF533941_1.txt,
 PCV1num48gbEF493843_1.txt, PCV1num49gbDQ648032_1.txt, PCV1num4gbKJ746929_1.txt,
 PCV1num50gbDQ358813_1.txt, PCV1num5gbKJ408799_1.txt, PCV1num6gbKJ408798_1.txt,
 PCV1num7gbAY754015_1.txt, PCV1num8gbAY754014_1.txt, PCV1num9gbKF732857_1.txt}
```

This mixed up the order a bit, but doesn't really matter what order PCV1ListofLists is in since we can just make a new list with these more well-ordered

```
PCV1ListofLists[[2]]
Length[PCV1ListofLists]
```

```
PCV1num11gbKC894933_1.txt
```

```
50
```

```
ReadList[StringJoin[PCV1directory, "PCV1num11gbKC894933_1.txt"]];
```

```
ReadList[StringJoin[PCV1directory, PCV1ListofLists[[2]]]];
```

```
PCV1Individual_1 = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[1]]]];
```

```
Do[PCV1Individual_k =
   ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k]]]], {k, 1, 7}]
Do[PCV1Individual_k = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k + 1]]]],
  {k, 8, 49}]
PCV1Samples = Table[PCV1Individual_k, {k, 1, 49}];
(*This is a list of the PCV1 samples with the differently-lengthed one,
 PCV1ListofLists[[8]], excised *)
```

```
PCV1Individual_39 - PCV1Individual_49
 (*Can see a few differences between these two samples*)
```

```
{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -2,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, -1, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, -2, 0, 0, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0}
```

```
Length[PCV1ListofLists[[8]]]] = 1758 ≠ 1759 = length of the rest
   (*Excise this one just to be safe, since it's the only different one*)
```

```
Length[PCV1Samples]
Table[Length[PCV1Samples[[k]]], {k, 1, 49}]
Union[Table[Length[PCV1Samples[[k]]], {k, 1, 49}]]
Print["If the samples are of the same
   length, this union function should give only a single number"]
```

```
49
```

```
{1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759,
 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759,
 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759,
 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759, 1759}
```

```
{1759}
```

```
If the samples are of the same length, this union function should give only a single number
```

Form for Python

```
PCV1directory = "C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals\\";
PCV1ListofLists =
  Import["C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals"];
Do[PCV1Individual_k = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k]]]],
 {k, 1, 7}]
Do[PCV1Individual_k = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k + 1]]]],
 {k, 8, 49}]
PCV1Samples = Table[PCV1Individual_k, {k, 1, 49}];
(*This is a list of the PCV1 samples with the differently-lengthed one,
 PCV1ListofLists[[8]], excised *)
```

```
StringJoin[{PCV1pythonformdirectory, "PCV1_Sample_", ToString[k], ".txt"}]
```

```
C:\Users\George\Documents\SVD DNA stuff\PCV1SamplesPythonForm\PCV1_Sample_k.txt
```

```
(*PCV1pythonformdirectory=
 "C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1SamplesPythonForm\\";
Do[Export[StringJoin[{PCV1pythonformdirectory,"PCV1_Sample_",ToString[k],".txt"}],
  PCV1Samples[[k]]],{k,1,49}]*)
```

Compressibility

Initial (some of which unusable) samples, See next section for proper samples

```
PCV1directory = "C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals\\";
PCV1ListofLists =
  Import["C:\\Users\\George\\Documents\\SVD DNA stuff\\PCV1 individuals"];
Do[PCV1Individual_k = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k]]]],
 {k, 1, 7}]
Do[PCV1Individual_k = ReadList[StringJoin[PCV1directory, PCV1ListofLists[[k + 1]]]]],
 {k, 8, 49}]
PCV1Samples = Table[PCV1Individual_k, {k, 1, 49}];
(*This is a list of the PCV1 samples with the differently-lengthed one,
 PCV1ListofLists[[8]], excised *)
```

The above blue cell defines a list of lists PCV1Samples. Each PCV1Samples[[k]] fo k from 1 to 49 corresponds to the complete genome of a PCV1 individual listed in the folder *PCV1 individuals*.

NOTE : The name of these text files have nothing to do with the order of PCV1Samples, since the order in the folder was scrambled when importing it to mathematica

```
Length[PCV1Samples]
```

```
49
```

```
PCV1Samples[[1]]
```

```
{0, 1, 1, 0, 2, 1, 2, 1, 0, 1, 3, 3, 1, 2, 2, 1, 0, 2, 1, 2, 2, 1, 0, 2, 1, 0, 1, 1, 3, 1, 2, 2, 1, 0,
 2, 1, 2, 3, 1, 0, 2, 3, 2, 0, 0, 0, 0, 3, 2, 1, 1, 0, 0, 2, 1, 0, 0, 2, 0, 0, 0, 0, 2, 1, 2, 2,
 1, 1, 1, 2, 1, 0, 0, 1, 1, 1, 1, 0, 3, 0, 0, 2, 0, 2, 2, 3, 2, 2, 2, 3, 2, 3, 3, 1, 0, 1, 1, 1,
 3, 3, 0, 0, 3, 0, 0, 3, 1, 1, 3, 3, 1, 1, 2, 0, 2, 2, 0, 2, 2, 0, 2, 0, 0, 0, 0, 0, 1, 0, 0, 0,
 0, 3, 0, 1, 2, 2, 2, 0, 2, 1, 3, 3, 1, 1, 0, 0, 3, 1, 3, 1, 1, 1, 3, 3, 3, 3, 3, 2, 0, 3, 3, 0,
 3, 3, 3, 3, 2, 3, 3, 3, 2, 1, 2, 2, 0, 2, 0, 2, 2, 0, 0, 2, 2, 3, 3, 3, 2, 2, 0, 0, 2, 0, 2, 2,
 2, 3, 0, 2, 0, 0, 1, 3, 1, 1, 3, 1, 0, 1, 1, 3, 1, 1, 0, 2, 2, 2, 2, 3, 3, 3, 2, 1, 2, 0, 0, 3,
 3, 3, 3, 2, 1, 3, 0, 0, 2, 0, 0, 2, 1, 0, 2, 0, 1, 3, 3, 3, 3, 0, 0, 1, 0, 0, 2, 2, 3, 2, 0, 0,
 2, 3, 2, 2, 3, 0, 3, 3, 3, 3, 2, 2, 3, 2, 1, 1, 1, 2, 1, 3, 2, 1, 1, 0, 1, 0, 3, 1, 2, 0, 2, 0,
 0, 0, 2, 1, 2, 0, 0, 0, 2, 2, 0, 0, 1, 1, 2, 0, 1, 1, 0, 2, 1, 0, 2, 0, 0, 3, 0, 0, 0, 2, 0,
 0, 3, 0, 1, 3, 2, 1, 0, 2, 3, 0, 0, 0, 2, 0, 0, 2, 2, 1, 1, 0, 1, 0, 3, 0, 1, 3, 3, 0, 3, 1,
 2, 0, 2, 3, 2, 3, 2, 2, 0, 2, 1, 3, 1, 1, 2, 1, 2, 2, 0, 0, 1, 1, 0, 2, 2, 2, 2, 0, 0, 2, 1,
 2, 1, 0, 2, 1, 2, 0, 1, 1, 3, 2, 3, 1, 3, 0, 1, 3, 2, 1, 3, 2, 3, 2, 0, 2, 3, 0, 1, 1, 1, 3,
 3, 3, 3, 2, 2, 0, 2, 0, 1, 2, 2, 2, 2, 3, 1, 3, 3, 3, 2, 2, 3, 2, 0, 1, 3, 2, 3, 0, 2, 1, 1,
 2, 0, 2, 1, 0, 2, 3, 3, 1, 1, 1, 3, 2, 3, 0, 0, 1, 2, 3, 3, 3, 2, 3, 1, 0, 2, 0, 0, 0, 3, 3,
 3, 1, 1, 2, 1, 2, 2, 2, 1, 3, 2, 2, 1, 3, 2, 0, 0, 1, 3, 3, 3, 3, 2, 0, 0, 0, 2, 3, 2, 0, 2,
 1, 2, 2, 2, 0, 0, 2, 0, 3, 2, 1, 0, 2, 1, 0, 2, 1, 2, 3, 2, 0, 3, 3, 2, 2, 0, 0, 2, 0, 1, 0,
 2, 1, 3, 2, 3, 0, 1, 0, 1, 2, 3, 1, 0, 3, 0, 2, 3, 2, 2, 2, 1, 1, 1, 2, 1, 1, 1, 2, 2, 3, 3,
 2, 3, 2, 2, 2, 0, 0, 2, 0, 2, 1, 1, 0, 2, 3, 2, 2, 2, 1, 1, 1, 2, 3, 0, 0, 3, 3, 3, 3, 2, 1, 3,
 2, 0, 2, 1, 1, 3, 0, 2, 1, 2, 0, 1, 0, 1, 1, 3, 0, 1, 3, 2, 2, 0, 0, 2, 1, 1, 3, 0, 2, 3, 0, 2,
 0, 0, 0, 3, 0, 0, 2, 3, 2, 2, 3, 2, 2, 2, 0, 3, 2, 2, 0, 3, 0, 3, 1, 0, 0, 2, 2, 0, 2, 0, 0,
 2, 0, 0, 2, 3, 3, 2, 3, 3, 2, 3, 3, 3, 3, 2, 2, 0, 3, 2, 0, 3, 3, 3, 3, 3, 0, 3, 2, 2, 1, 3, 2,
 2, 3, 3, 0, 1, 1, 3, 3, 2, 2, 2, 0, 3, 2, 0, 3, 1, 3, 0, 1, 3, 2, 0, 2, 0, 1, 3, 2, 3, 2, 3, 2,
 0, 1, 1, 2, 2, 3, 0, 3, 1, 1, 0, 3, 3, 2, 0, 1, 3, 2, 3, 0, 2, 0, 2, 0, 1, 3, 0, 0, 0, 2, 2, 2,
 2, 2, 3, 0, 1, 3, 2, 3, 3, 1, 1, 3, 3, 3, 3, 3, 3, 2, 2, 1, 1, 1, 2, 1, 0, 2, 3, 0, 3, 3, 3, 3,
 2, 0, 3, 3, 0, 1, 1, 0, 2, 1, 0, 0, 3, 1, 0, 2, 2, 1, 1, 1, 1, 1, 1, 0, 2, 2, 0, 0, 3, 2, 2, 3,
 0, 1, 3, 1, 1, 3, 1, 0, 0, 1, 3, 2, 1, 3, 2, 3, 1, 1, 1, 0, 2, 1, 3, 2, 3, 0, 2, 0, 0, 2, 1,
 3, 1, 3, 1, 3, 0, 3, 1, 2, 2, 0, 2, 2, 0, 3, 3, 0, 1, 3, 0, 1, 3, 3, 3, 2, 1, 0, 0, 3, 3, 3, 3,
 2, 2, 0, 0, 2, 0, 1, 3, 2, 1, 3, 2, 2, 0, 2, 0, 0, 1, 0, 0, 3, 1, 1, 0, 1, 2, 2, 0, 2, 2, 3, 0,
 1, 1, 1, 2, 0, 0, 2, 2, 1, 1, 2, 0, 3, 3, 3, 2, 0, 0, 2, 1, 0, 2, 3, 2, 2, 0, 1, 1, 1, 0, 1, 1,
 1, 3, 2, 3, 2, 1, 1, 1, 3, 3, 3, 3, 1, 1, 1, 0, 3, 0, 3, 0, 0, 0, 0, 3, 0, 0, 0, 3, 3, 0, 1, 3,
 2, 0, 2, 3, 1, 3, 3, 3, 3, 3, 3, 2, 3, 3, 0, 3, 1, 0, 1, 0, 3, 1, 2, 3, 0, 0, 3, 2, 2, 3, 3,
 3, 3, 3, 0, 3, 3, 3, 3, 3, 0, 3, 3, 3, 0, 3, 3, 3, 2, 2, 0, 2, 2, 2, 3, 1, 3, 3, 3, 3, 0, 2, 2,
 0, 3, 0, 0, 0, 3, 3, 1, 3, 1, 3, 2, 0, 0, 3, 3, 2, 3, 0, 1, 0, 3, 0, 0, 0, 3, 0, 2, 3, 1, 0,
 2, 1, 1, 3, 3, 0, 1, 1, 0, 1, 0, 3, 0, 0, 3, 3, 3, 3, 2, 2, 2, 1, 3, 2, 3, 2, 2, 1, 3, 2, 1,
 0, 3, 3, 3, 3, 2, 2, 0, 2, 1, 2, 1, 0, 3, 0, 2, 1, 1, 2, 0, 2, 2, 1, 1, 3, 2, 3, 2, 3, 2, 1,
 3, 1, 2, 0, 1, 0, 3, 3, 2, 2, 3, 2, 3, 2, 2, 2, 3, 0, 3, 3, 3, 0, 0, 0, 3, 2, 2, 0, 2, 1, 1,
 0, 1, 0, 2, 1, 3, 2, 2, 3, 3, 3, 1, 3, 3, 3, 3, 0, 3, 3, 0, 3, 3, 3, 2, 2, 2, 3, 2, 2, 0, 0,
 1, 1, 0, 0, 3, 1, 0, 0, 3, 3, 2, 3, 3, 3, 2, 2, 3, 1, 1, 0, 2, 1, 3, 1, 0, 2, 2, 3, 3, 3, 2,
 2, 2, 2, 2, 3, 2, 0, 0, 2, 3, 0, 1, 1, 3, 2, 2, 0, 2, 3, 2, 2, 3, 0, 2, 2, 3, 0, 0, 0, 2, 2,
 2, 1, 3, 2, 1, 1, 3, 3, 0, 3, 2, 2, 3, 2, 3, 2, 2, 1, 2, 2, 2, 0, 2, 2, 0, 2, 3, 0, 2, 3, 3,
 0, 0, 3, 0, 3, 0, 2, 2, 2, 2, 3, 1, 0, 3, 0, 2, 2, 1, 1, 0, 0, 2, 3, 3, 2, 2, 1, 2, 2, 0, 2,
 2, 2, 2, 2, 3, 3, 0, 1, 0, 0, 0, 2, 3, 3, 2, 2, 1, 0, 3, 1, 1, 0, 0, 2, 0, 3, 0, 0, 1, 0, 0,
 1, 0, 2, 3, 2, 2, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 3, 1, 3, 3, 3, 2, 0, 3, 3, 0, 2, 0, 2, 2, 3,
 2, 0, 3, 2, 2, 2, 2, 3, 1, 3, 1, 3, 3, 2, 2, 2, 3, 0, 0, 0, 0, 3, 3, 1, 0, 3, 0, 3, 3, 3, 0,
 2, 1, 1, 3, 3, 3, 1, 3, 0, 0, 3, 0, 1, 2, 2, 3, 0, 2, 3, 0, 3, 3, 2, 2, 0, 0, 0, 2, 2, 3, 0,
 2, 2, 2, 2, 3, 0, 2, 2, 2, 2, 2, 3, 3, 2, 2, 3, 2, 1, 1, 2, 1, 1, 3, 2, 0, 2, 2, 2, 2, 2, 2,
 2, 0, 2, 2, 0, 0, 1, 3, 2, 2, 1, 1, 2, 0, 3, 2, 3, 3, 2, 0, 0, 3, 3, 3, 2, 0, 2, 2, 3, 0, 2,
 3, 3, 0, 0, 1, 0, 3, 3, 1, 1, 0, 0, 2, 0, 3, 2, 2, 1, 3, 2, 1, 2, 0, 2, 3, 0, 3, 1, 1, 3, 1,
 1, 3, 3, 3, 3, 0, 3, 2, 2, 3, 2, 0, 2, 3, 0, 1, 0, 0, 0, 3, 3, 1, 3, 2, 3, 0, 2, 0, 0, 0, 2,
 2, 1, 2, 2, 2, 0, 0, 3, 3, 2, 0, 0, 2, 2, 3, 0, 1, 1, 1, 2, 3, 1, 3, 3, 3, 1, 2, 2, 1, 2, 1,
 1, 0, 3, 1, 3, 2, 3, 0, 0, 1, 2, 2, 3, 3, 3, 1, 3, 2, 0, 0, 2, 2, 1, 2, 2, 2, 2, 3, 2, 3, 2,
 1, 1, 0, 0, 0, 3, 0, 3, 2, 2, 3, 1, 3, 3, 1, 3, 1, 1, 2, 2, 0, 2, 2, 0, 3, 2, 3, 3, 3, 1, 1,
 0, 0, 2, 2, 3, 2, 2, 1, 3, 2, 1, 2, 2, 2, 2, 2, 1, 2, 2, 2, 3, 1, 1, 3, 3, 1, 3, 3, 1, 3, 2,
 1, 2, 2, 3, 0, 0, 1, 2, 1, 1, 3, 1, 1, 3, 3, 2, 2, 1, 1, 0, 1, 2, 3, 1, 0, 3, 1, 1, 3, 0, 3,
 0, 0, 0, 0, 2, 3, 2, 0, 0, 0, 2, 0, 0, 2, 3, 2, 1, 2, 1, 3, 2, 1, 3, 2, 3, 0, 2, 3, 2, 3, 3}
```

Comparison of sample ordering

```
testset = {0, 1, 0, 2, 0, 3, 0, 0, -1, -2, -1, -2, 0, 0, -3, 0};
Count[testset, 0]
Length[testset]
Length[testset] - Count[testset, 0] (*Number of nonzero elements *)
```

```
8
```

```
16
```

```
8
```

```
indivcomparison[k_] := PCV1Individual_1 - PCV1Individual_k
numofdiffelements[k_] := Length[indivcomparison[k]] - Count[indivcomparison[k], 0]
  (*Number of different elements between PCV1Individual_1 and PCV1Individual_k *)
```

```
Mean[Table[numofdiffelements[k], {k, 1, 49}]] // N
Median[Table[numofdiffelements[k], {k, 1, 49}]] // N
```

```
229.878
```

```
16.
```

```
Do[If[numofdiffelements[k] > 30, Print[k], 0], {k, 1, 49}]
(*Gives us the k for which the number of different elements is more than 30 *)
```

```
7

11

12

21

23

29

33

40

43

47

48
```

```
(*Do[Print[{k,numofdiffelements[k]}],{k,1,49}] (*To check the above*)*)
```

Note that there are 11 samples for which the number of different elements is more than 30, delete these in the python code and rename some others as follows:

Now we remove these and redefine for k of PCV1_Sample_k.txt

k= 49 -> 7
46 to 11
45 to 12
44 to 21
42 to 23
41 to 29
39 to 33

So we now have 49 - 11 = 38 usable samples

Usable Samples

```
PCV1Sample[k_] := ReadList[StringJoin[
   "C:\\Users\\George\\Documents\\SVD DNA stuff\\Usable PCV1 Samples\\PCV1_Sample_",
   ToString[k], ".txt"]]

(*This is a defines PCV1Sample[k] as the list form of PCV1_Sample_k.txt ,
the k-th Usable PCV1 samples,
THE NUMBERS DONT ALL LINE UP WITH THOSE IN PREVIOUS PARTS,
I.E. PCV1_Sample_38.txt  here is NOT same as the old PCV1_Sample_38.txt,
renumbered them *)
```

```
Length[(PCV1Sample[1] - PCV1Sample[9])] -
 Count[(PCV1Sample[1] - PCV1Sample[2]), 0] (*Number of differences *)
```

```
30
```

```
(PCV1Sample[1] - PCV1Sample[9])[[k]]
```

```
{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0,
  -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, -3, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, -2, 0, 0, -1, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -1, 0, -2, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, -2, 0, 0, 0, 0,
  0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0}⟦k⟧
```

```
Union[Table[If[(PCV1Sample[1] - PCV1Sample[9])[[k]] == 0, 0, k], {k, 1, 1759}]]
(*Places where these samples differ*)
```

```
{0, 465, 469, 658, 760, 907, 910, 916, 956, 958, 1032, 1216,
 1321, 1346, 1378, 1400, 1503, 1510, 1519, 1578, 1587, 1670, 1757}
```

```
differingbasepairlocations[i_, j_] :=
 Union[Table[If[(PCV1Sample[i] - PCV1Sample[j])[[k]] == 0, 0, k], {k, 1, 1759}]]
 (*Places where these samples differ*)
```

```
Table[differingbasepairlocations[i, 9], {i, 1, 8}]
```

```
$Aborted
```