

```
In [4]: # Load and process the protein sequences from the uploaded file
file_path = "K13_Protein_genes.txt"

# Read the file content
with open(file_path, "r") as file:
    data = file.readlines()
```

```
In [5]: # Prints the entire file as a single string to validate the protein sequence
print("Sequence List for Mutation Alayisis:")
print("".join(data))
print("\n" + "="*50 + "\n")
```

Sequence List for Mutation Alalysis:

>K13_Target1

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKKFSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK
LFKDKKELEMERVRLYKELENRKNIEEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFDDEERLRLQEIDKIKLVLYLEK
EKYYQYKKNFENDKKKIVDANIATETMIDINVGGAI FETSRLTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPISTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNFVATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target2

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKT FSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK
LFKDKKELEMERVRLYKELENRKNIEEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFDDEERLRLQEIDKIKLVLYLEK
EKYYQYKKNFENDKKKIVDANIATETMIDINVGGAI FETSRLTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNF GAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target3

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKT FSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK
LFKDKKELEMERVRLYKELENRKNIEEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFDDEERLRLQEIDKIKLVLYLEK
EKYYQYKKNFENDKKKIVDANIATETMIDINVGGAI FETSRLTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGTIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTLRSSAMTVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNF GAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target4

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKT FSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK
LFKDKKELEMERVRLYKELENRKNIEEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFDDEERLRLQEIDKIKLVLYLEK
EKYYQYKKNFENDKKKIVDANIATETMIDINVGGAI FETSRLTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNF GAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target5

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKT FSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK
LFKDKKELEMERVRLYKELENRKNIEEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFDDEERLRLQEIDKIKLVLYLEK
EKYYQYKKNFENDKKKIVDANIATETMIDINVGGAI FETSRLTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNF GAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target6

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYLNLKLNKKKDTITNENNNINHNHNNNNNL TANNITNNLINNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKKFSTVNNVNDTYEKKIIETELSDASDFENMVGDLRITFINWLKKTQMNFIREDK

LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMTVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNFGAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target7

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYL NKL LNKKKDTITNENNNINHNHNNNNNL TANNITNNL INNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKTFTSVNNVNDTYEKKI IETELSDASDFENMVGD LRITFINWLKKTQ MNFIREKDK
LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNFGAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target8

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYL NKL LNKKKDTITNENNNINHNHNNNNNL TANNITNNL INNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKTFTSVNNVNDTYEKKI IETELSDASDFENMVGD LRITFINWLKKTQ MNFIREKDK
LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEE VAPLNTPRSSAMCVAFDNKIYVIGGTNGER
LNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILDSVEQYQPFNKRWQFLNGVPEKKMNFGAATL
SDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target9

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYL NKL LNKKKDTITNENNNINHNHNNNNNL TANNITNNL INNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKTFTSVNNVNDTYEKKI IETELSDASDFENMVGD LRITFINWLKKTQ MNFIREKDK
LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNFGAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target10

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYL NKL LNKKKDTITNENNNINHNHNNNNNL TANNITNNL INNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKTFTSVNNVNDTYEKKI IETELSDASDFENMVGD LRITFINWLKKTQ MNFIREKDK
LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLYVFGGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMCVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILD
SVEQYQPFNKRWQFLNGVPEKKMNFGAATLSDSYIITGGENG EVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVL IANI

>K13_Target11

MEGEKVTKANSISNFSMTYDRESGGNSNSDDKSGSSSEENDSNSFMNLTSDKNEKTENNSFLLNNSSYGNVKDSLLESIDM
SVLDSNFDSKKDFLPSNLSRTFNMSKDNIGNKYL NKL LNKKKDTITNENNNINHNHNNNNNL TANNITNNL INNNMNSPSI
MNTNKKENFLDAANLINDDSGLNNLKKFTSVNNVNDTYEKKI IETELSDASDFENMVGD LRITFINWLKKTQ MNFIREKDK
LFKDKKELEMERVRLYKELENRKNIIEQKLHDERKKLDIDISNGYKQIKKEKEEHRKRFD EERLRF LQEIDKIKLVLYLEK
EKYYQEYKNFENDKKKIVDANIATETMIDINVGGAI FETS RHTLTQQKDSFIEKLLSGRHHVTRDKQGRIFLDRDSEL FRI
ILNFLRNPLTIPIPKDLSESEALLKEAEFYGIKFLPFPLVFCIGGFDGVEYLN SMELLDISQQCWRMCTPMSTKKAYFGSA
VLNNFLHVF GGNNYDYKALFETEVDRLRDVWYVSSNLNIPRRNNCGVTSNGRIYCIGGYDGSSIIPNVEAYDHRMKAWVE
VAPLNTPRSSAMTVAFDNKIYVIGGTNGERLNSIEVYEEKMNKWEQFPYALLEARSSGA AFNYLNQIYVVGGIDNEHNILD

SVEQYQPFNKRWQFLNGVPEKKMNFGAATLSDSYIITGGENGVLNSCHFFSPDTNEWQLGPSLLVPRFGHSVLIANI

=====

```
In [6]: from tabulate import tabulate
# Parse the FASTA file format
sequences = {}
current_target = None

for line in data:
    line = line.strip()
    if line.startswith(">"): # New target identifier
        current_target = line[1:].strip()
        sequences[current_target] = ""
    else:
        sequences[current_target] += line # Append protein sequence

# Known resistance mutations and their reference amino acids
known_mutations = {
    580: "C", # C580Y (Cysteine → Tyrosine)
    539: "R", # R539T (Arginine → Threonine)
    493: "Y", # Y493H (Tyrosine → Histidine)
    675: "A", # A675V (Alanine → Valine)
    574: "P", # P574L (Proline → Leucine)
    476: "M", # M476I (Methionine → Isoleucine)
}

# Identify mutations in each target sequence
detected_mutations = []

for target, sequence in sequences.items():
    for pos, ref_aa in known_mutations.items():
        if pos <= len(sequence):
            observed_aa = sequence[pos - 1]
            if observed_aa != ref_aa:
                detected_mutations.append([target, pos, ref_aa, observed_aa])

# Print detected mutations in a table format
if detected_mutations:
    print(tabulate(detected_mutations, headers=["Target", "Position", "Reference", "Observed"]))
else:
    print("No mutations detected.")
```

Target	Position	Reference Amino-Acid	Observed Amino-Acid
K13_Target1	675	A	V
K13_Target1	476	M	I
K13_Target3	580	C	T
K13_Target3	539	R	T
K13_Target3	574	P	L
K13_Target6	580	C	T
K13_Target8	580	C	E
K13_Target8	539	R	V
K13_Target8	574	P	T
K13_Target8	476	M	I
K13_Target11	580	C	T
K13_Target11	493	Y	H