

Supplementary Figure 1

```
In [32]: ## Importing pandas to read .csv data
import pandas as pd
import numpy as np

## uploading data as variable df
df = pd.read_excel("AGR_RWPE_PTN2/RWPE1_PTN2_opensearch/RWPE1_PTN2_opensearch/RWPE1_2/RWPE1_2_peptide.xlsx")
xref_df = pd.read_excel("Cross_Referencing_Databases/Xref_master.xlsx")
print(df.shape)

(14349, 17)
```

```
In [33]: ##remove any empty columns
df.dropna(how='all', axis=1, inplace=True)

#display(df)
```

```
In [34]: ## remove rows if protein is not reviewed
df = df[df['Status'].notna()]
```

```
In [35]: ## new dataframe created with proteins only in certain columns of Xref_master

df_relaxedsurfaceome = df[df['Gene'].isin(xref_df['Nix_surfaceome'])]
#display(df_relaxedsurfaceome)
df_stringentsurfaceome = df[df['Gene'].isin(xref_df['Wollscheid_surfaceome'])]
#display(df_stringentsurfaceome)
print(df_relaxedsurfaceome.shape)
print(df_stringentsurfaceome.shape)

(1478, 17)
(728, 17)
```

```
In [28]: ##Export to an excel file
df_relaxedsurfaceome.to_excel('Sample_Nix.xlsx')
df_stringentsurfaceome.to_excel('Sample_Wollscheid.xlsx')
```

```
In [ ]:
```