

scvdj report

```
import muon
import pandas as pd
import seaborn as sns
import os

PROJ="/home/ubuntu/projmnt/betts/coculture"

# mdata = muon.read_h5mu("/home/ubuntu/projmnt/betts/coculture/rna_adt_vdj_scvi_integrated.h5")
nt_annot = pd.read_csv(os.path.join(PROJ, "ir_query_annotate/levenshtein_nt_annot.csv"), index_col=0)
```

```
/tmp/ipykernel_16055/2758453826.py:9: DtypeWarning: Columns (21,27,35) have mixed types. Specify dtype option on import or setting low_resolution=False or infer_min_count=0 in
nt_annot = pd.read_csv(os.path.join(PROJ, "ir_query_annotate/levenshtein_nt_annot.csv"), index_col=0)
```

For a demonstration of a line plot on a polar axis, see [?@fig-polar](#).

```
ref_matches = nt_annot.sequence_id.value_counts()
ref_match_counts = pd.DataFrame(pd.cut(ref_matches, bins = [0,1,5,10,100,1000]).value_counts())
sns.barplot(data = ref_match_counts, x="bin", y = "count")
```

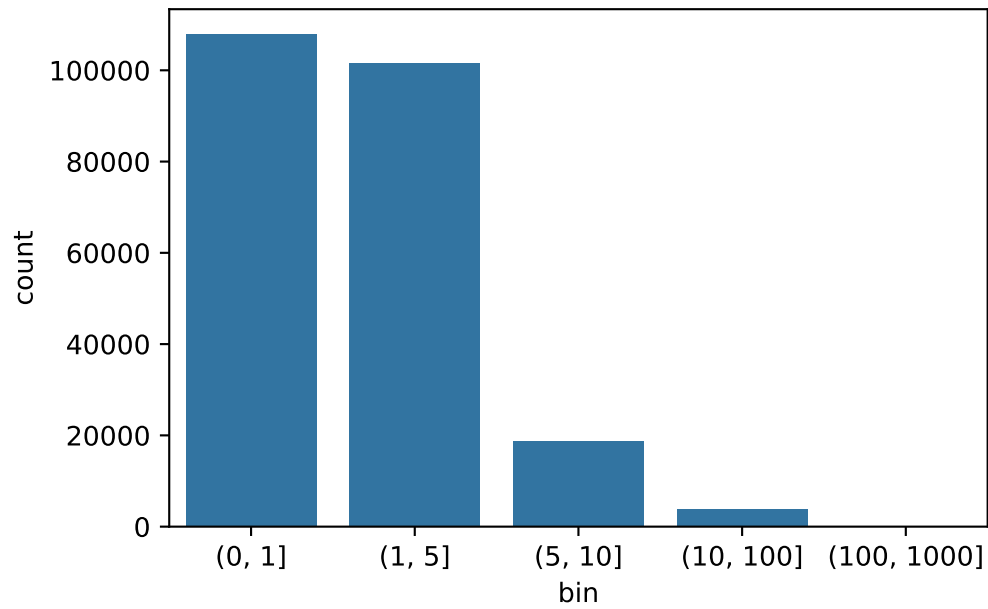


Figure 1: Bar plot showing the number of query junction nucleotide sequences that match those in the reference.

```
cell_matches = nt_annot.index.value_counts()
cell_match_counts = pd.DataFrame(pd.cut(cell_matches, bins = [0,1,5,10,100,1000]).value_counts())
cell_match_counts.bin
sns.barplot(data = cell_match_counts, x="bin", y = "count")
# print(mdata["airr"].shape)
```

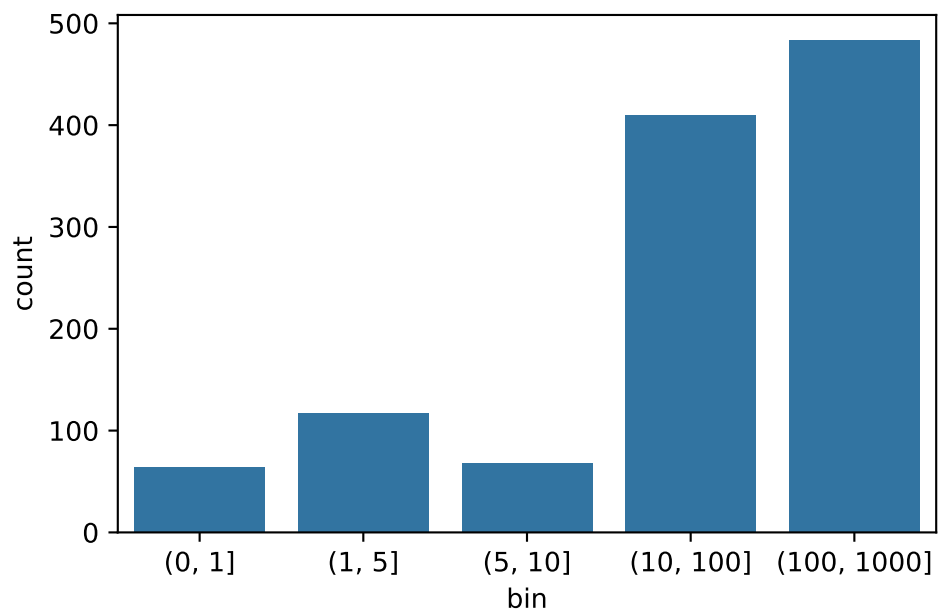


Figure 2: Bar plot showing the number of reference junction nucleotide sequences that are matched by thoes in the query.