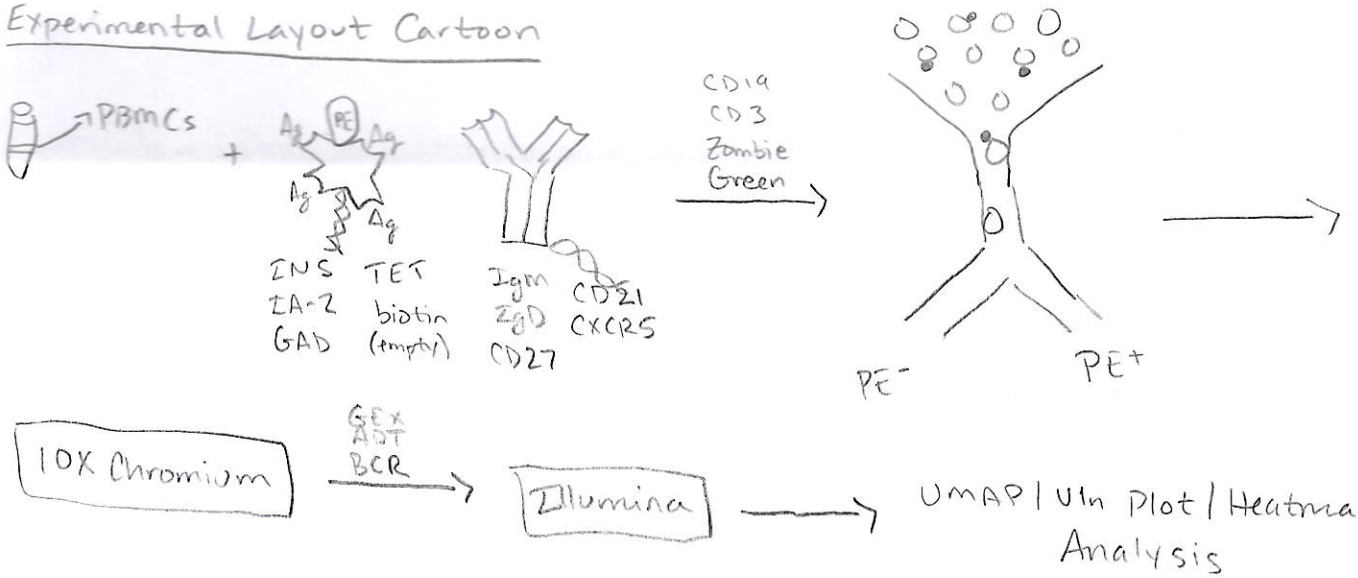
1) Donor Metadata

a) ID, Age*, Sex, Status*, Days Post Diagnosis, HLA, Autoantibodies, #cells captured

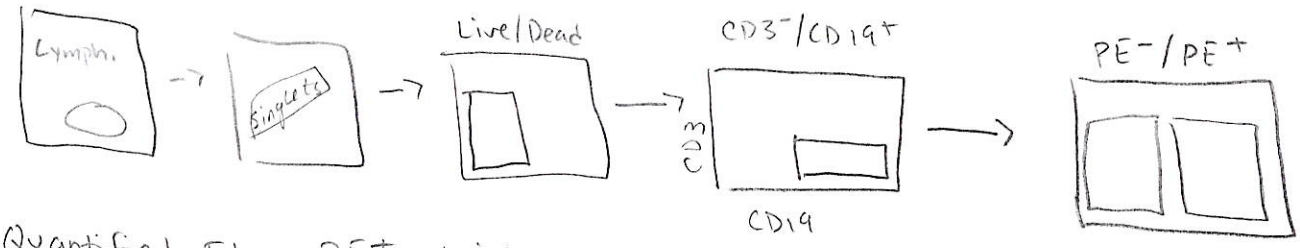
maybe make DR4 DQ8 IDs req throughout?

*table organized first by status, then by age w/in status

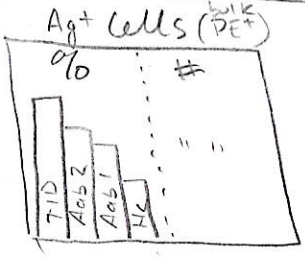
b) Experimental Layout Cartoon



c) Sort Gating Strategy



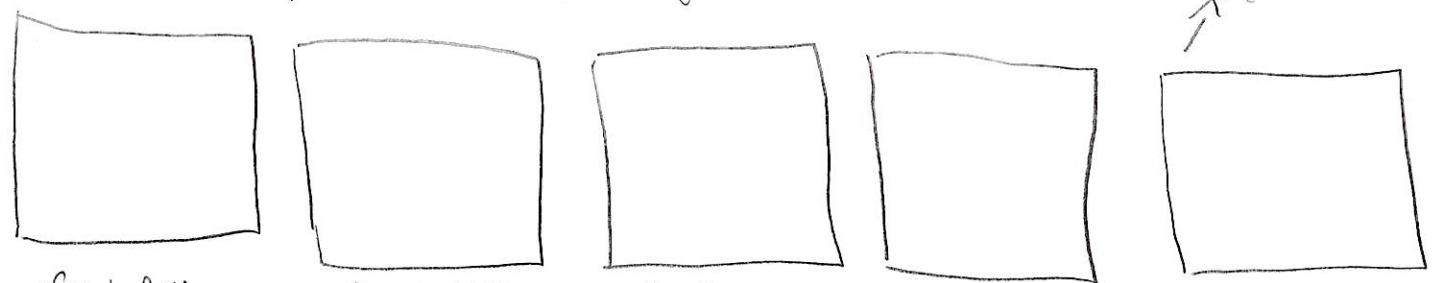
d) Quantified Flow PE+ staining



*Have individual Ag PE+ plots in supplemental figures

e) QC of sequencing data + cell type assignment

comment on PB identification (not doublets if <2 IgH)



SuPP

a) Table of donor metadata

b) Main steps of experiments

c) How cells were sorted

d) Ag^+ cells found by flow sorting as % total + absolute # by status

=> More PE^+ cells in TID (Aab?) than in HC

e) Seq. QC info

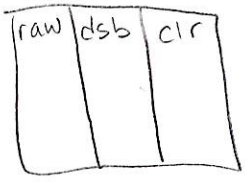
1. RNA count cutoff assignment
2. ADT count cutoff assignment
3. % mitochondrial reads cutoff assignment
4. Final merged integration UMAP (sample IDs)
5. Cell typing UMAP

=> Samples successfully integrated + define # B cell subtypes by RNA expression

(2) ADT analysis pipeline (a/b/c might be relegated to supplemental)

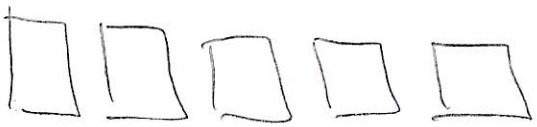
(Do we need any plot dis-proving INSR binding?)

a) dsb vs clr norm



heatmap

b) compare single Ag^+ flow plots



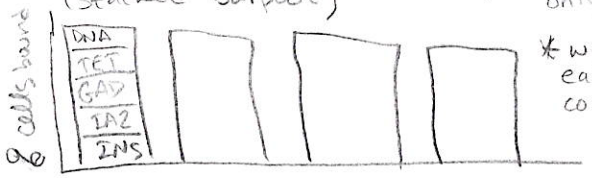
(look @ % + & compare to % + by seq?)

d) Clustering by individual ADT UMAPs

c) Barplot

"all type UMAPs w/ each Ag highlighted

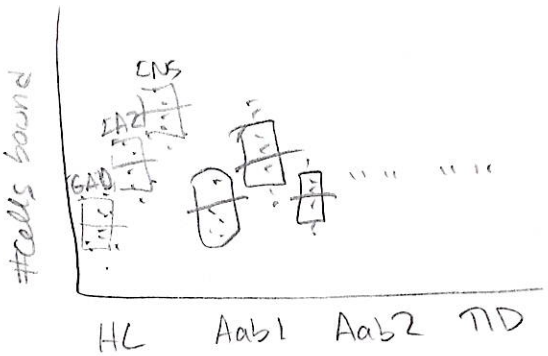
e) Quantify cells bound by Ag (mem vs naive only if interesting) (stacked barplot)



* write cell # in each bar, color code Antigens

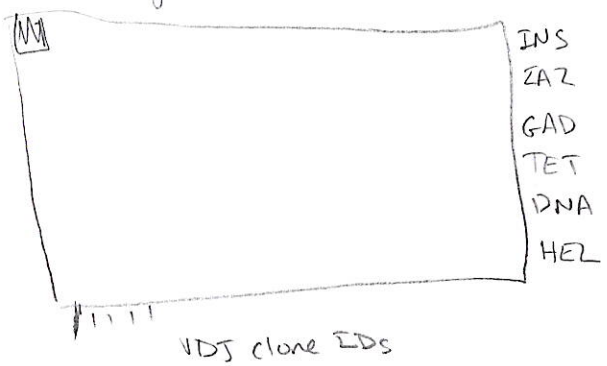
f) Correlate Ag binding by Known Aab status (are there more cells bound in individuals with known Abs detected?)

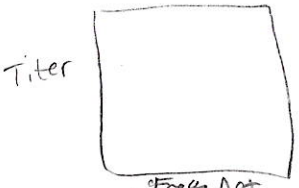
Dot Plot



*individual donors are color/shape/fill identifiable grouped by status + only denoted if known Aab^+, black circle otherwise (all HC = black circles)

g) BLI/ELISA affinity data here? (could also be attached to VDJ figure...) Heatmap





Heatmap

a) justification for using dsb over clr norm method

=> clr deviates from raw substantially, changing interpretation of data

b) single Ag stain flow

=> Dilution of Ag reagent used for each is comparable

=> All Ags found even in HC (as expected)

=> Background binding of empty (DNA) reagent is low

c) Ags bound across all B cell subtypes

d) Clustering by Ag binding identifies Ag-specific cells

=> aids in distinction between single & multi-binders

e) Quantify cells bound by Ag

=> more INS+ (IA2, GAD?) cells in T1D (Aab?) than in HC (significant?)

↳ if interesting, compare w/in B cell subtypes

f) Correlate Ag binding by known Aab status

=> more Ag+ cells in individuals w/ Aab against that Ag than no Aab donors

g) Ag/Ab affinity

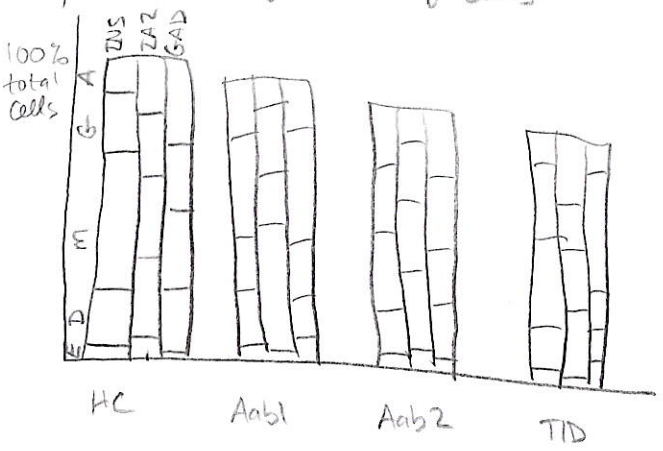
=> Higher affinity cells in T1D/Aab than HC?

↳ if not, discuss why method picks up low affinity cells

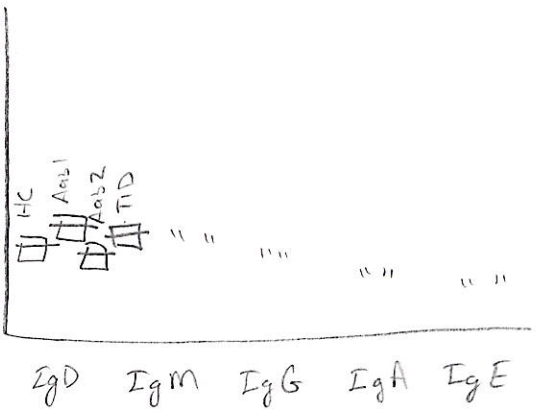
↳ maybe discuss regardless

(3) V(D)J Analysis

a) Isotype Frequency in Ag+ cells

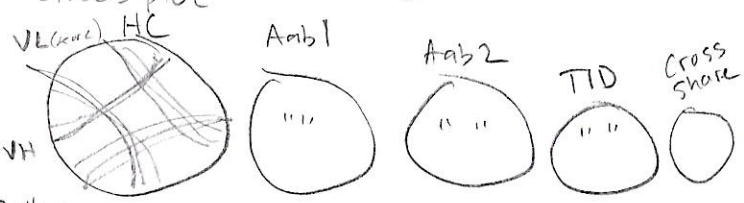


b) Frequency of mismatches/SHM



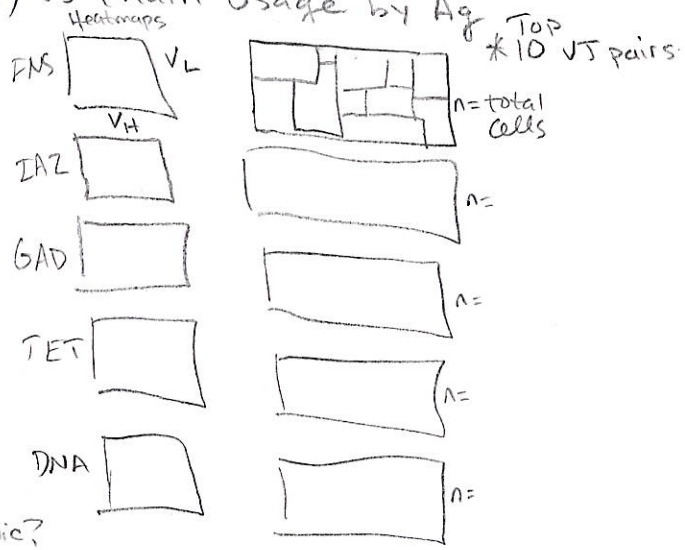
d) VJ chain usage sharing by status

circos plot



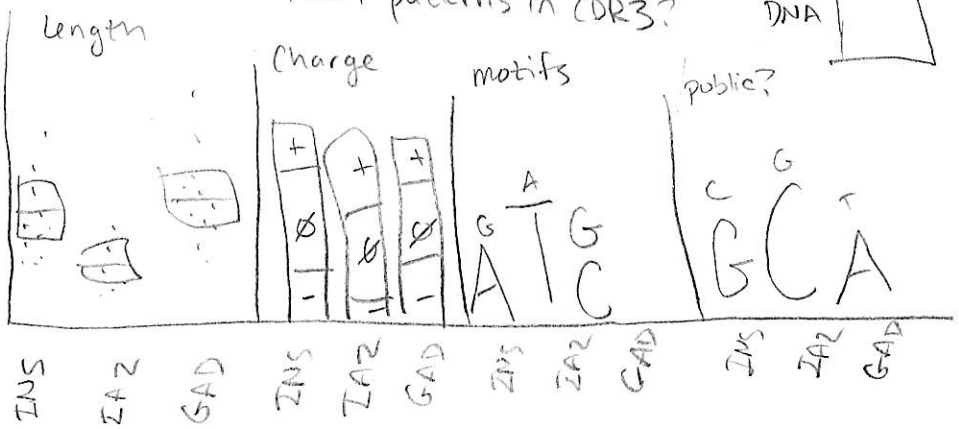
Top # H-L pairs shared in Ag+ cells (INS, IA2, GAD) by status

c) VJ chain usage by Ag



e) ?

In top/shared clones, any patterns in CDR3?



a) Isotype Frequency in Ag+ cells

=> mostly IgM -> reflects experiment's bias for picking up low affinity naive cells

=> maybe slightly more class-switched memory isotypes in T1D/Ab+, quantify if so

b) Frequency of mismatches/SHM

=> Generally unremarkable between groups -> b/c mostly naive

↳ what if looking only @ memory cells?

c) VJ chain usage by Ag

=> look @ Top 10 VJ pairs utilized for each Ag (maybe only INS/IA2/GAD)

d) VJ chain usage sharing by status

=> more shared VJ clones in T1D (Aab?) than HC

↳ also depict any cross-sharing between groups

e) In top 10 and/or shared clones, what patterns are interesting? (by Ag)
note status

=> CDR3 length

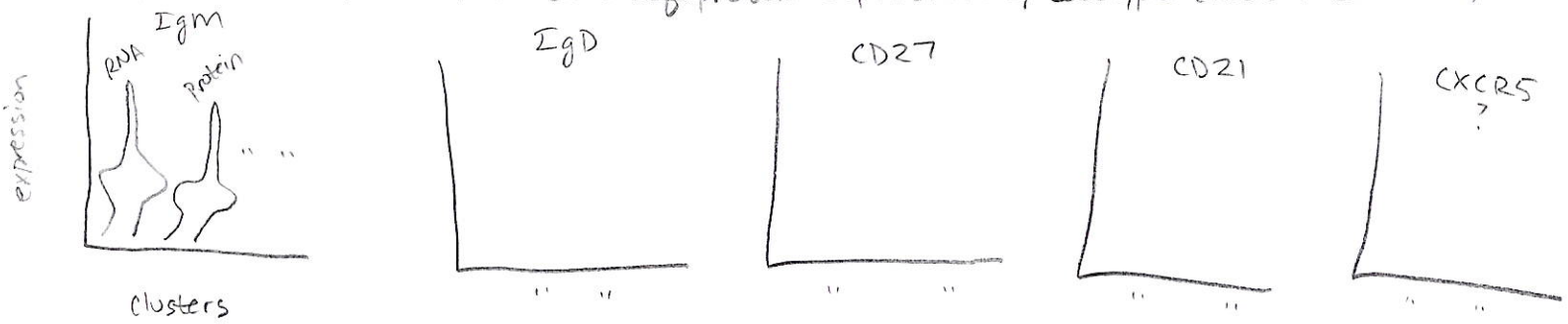
=> CDR3 charge

=> common CDR3 motifs

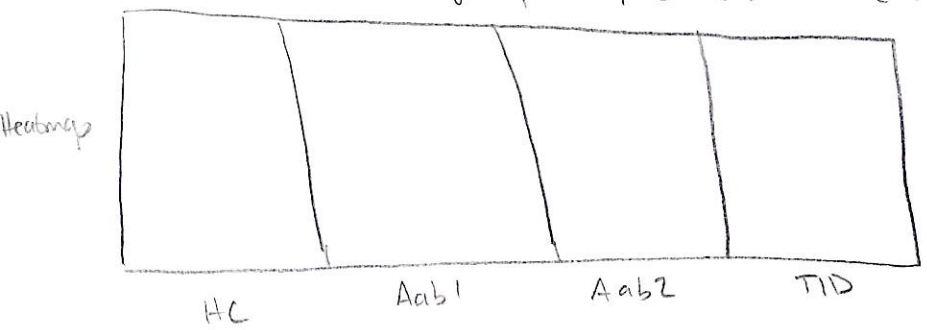
=> match any public clones?

④ GEX Analysis

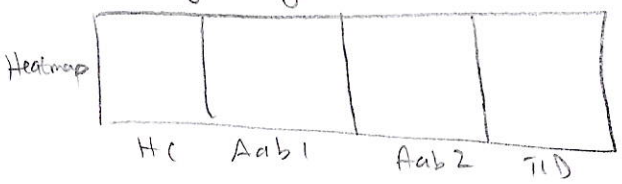
a) Compare RNA expression to CITE-seq protein expression by celltype cluster (Vln Plots)



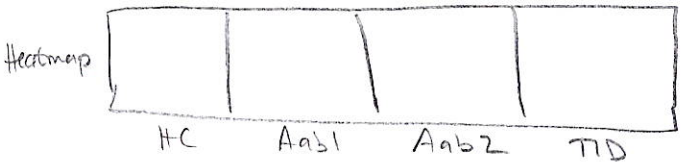
b) INS$^+$, IA2$^+$, GAD$^+$ cells grouped by status DE (*twins noted)



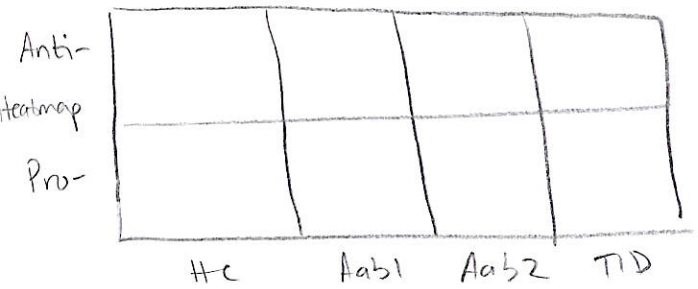
c) BCR signaling Genes



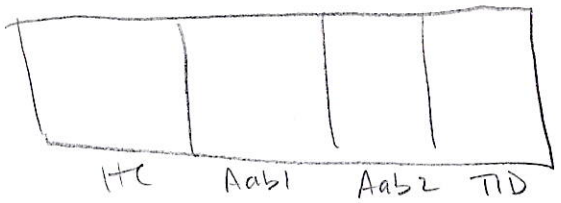
D) Ag processing + presentation Genes



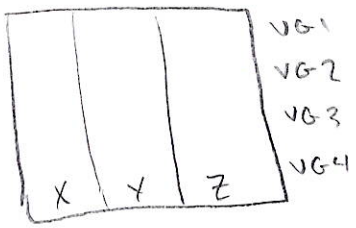
E) Cytokine signaling genes



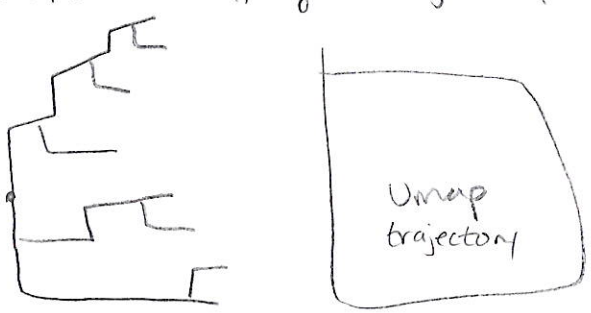
F) Viral Infection Response Genes?



G) Presence of EBV DNA?
Show any samples it is found in, and (noting status) which annotated viral genes are found



H)? BND/DN cell typing + trajectory analysis
(or by a)?



a) RNA expression by seq. & protein expresion by Ab-based detection seq. are not always in alignment, demonstrating the importance of looking @ both readouts in determining cell-type information

b) Full DE on Ag+ cells by status
   => unique signatures observed between groups

c) BCR Signaling
   => Up in TID (Aab?) compared to HC

d) Ag processing + presentation
   => Up in TID (Aab?) compared to HC

e) Cytokine Signaling
   => pro-inf Up in TID (Aab?) compared to HC
   => anti-inf up in HC compared to TID (Aab?)

f) Viral Response Genes
   => Up in TID (Aab?) compared to HC

g) Presence of EBV DNA?
   => maybe in TID ???

h) DN/BND cell typing trajectory towards PBs
   (maybe move up to B, closer to CITE-seq plots)
   => Look @ $INS^+/IA2^+/GAD^+$ cells RNA velocity
      from. naive -> BND -> DN -> PB
      ↳ are there differences by status?