Immunoinformatics
-
Computational approaches to study the human immune system

Katharina Imkeller
Goethe University Frankfurt

CSAMA 2026

What you will learn in this lecture

Get to know computational approaches to study:

1. Cell types and phenotypes
2. Interactions of immune cells
3. Antigen specificity

1. **Cell types and phenotypes**
2. Interactions of immune cells
3. Antigen specificity

How immunologists usually define cell types

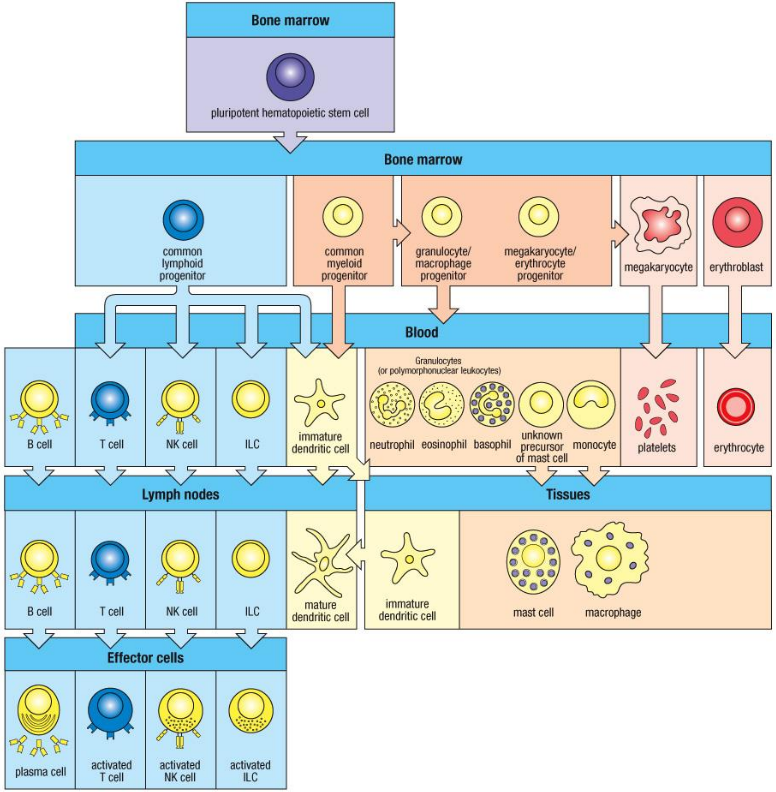
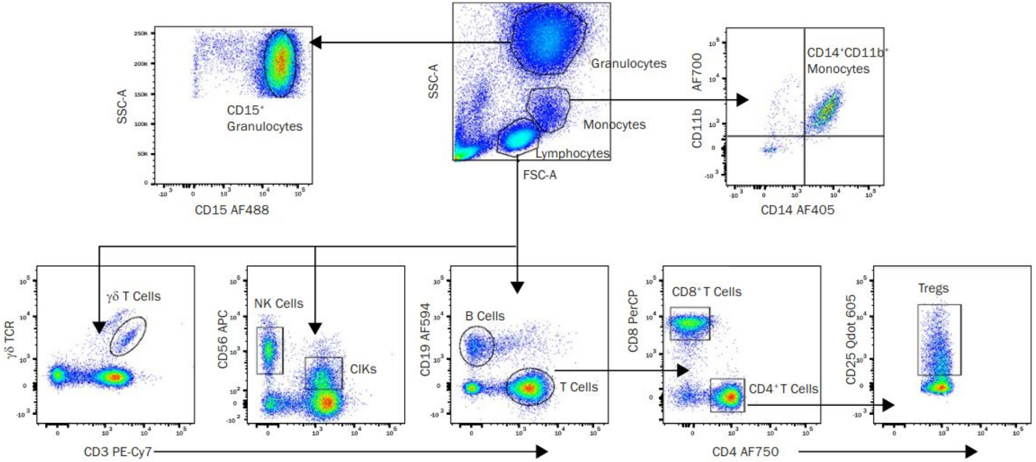


Figure 1.3 Janeway's Immunobiology, 9th ed. (© Garland Science 2017)

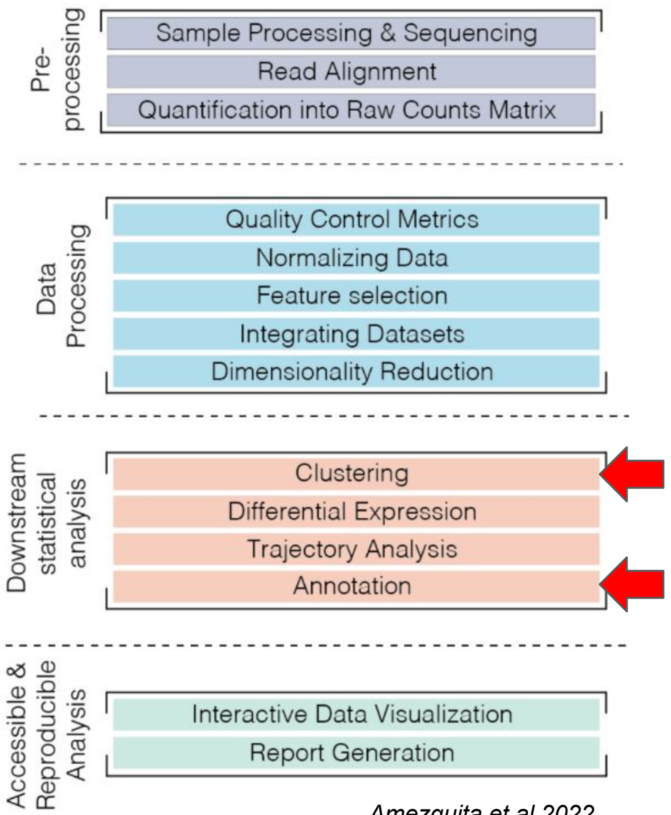
FACS gating for cell type identification



Cell type assignment in the single cell transcriptomics analysis workflow

Very likely:
scRNA cell type not exactly identical to FACS cell type

scSeq analysis workflow
(see single-cell lecture by D. Risso)

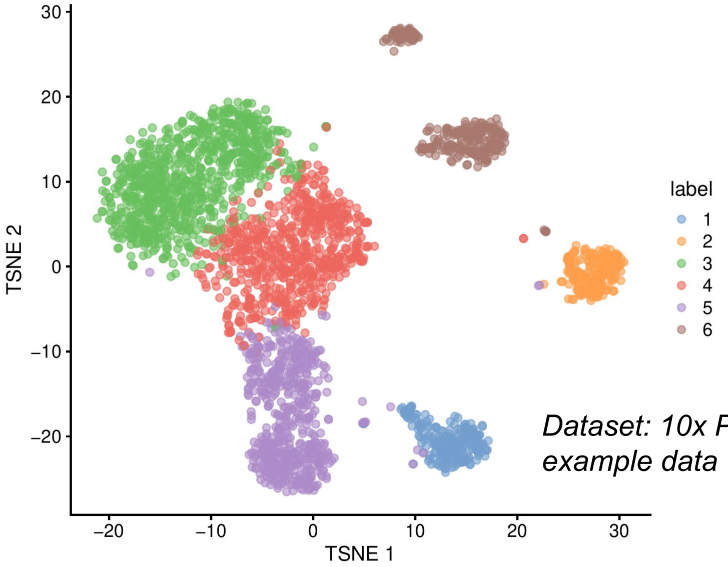


Amezquita et al 2022

Manual annotation using marker gene detection

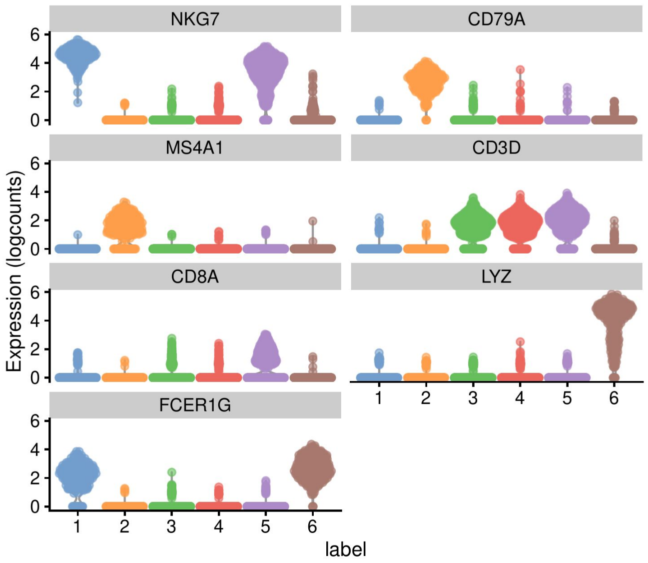
Bioconductor package: scran
Functions: scoreMarkers(), findMarkers()

T-SNE of cells coloured by cluster



Dataset: 10x PBMC example data

Expression of marker genes

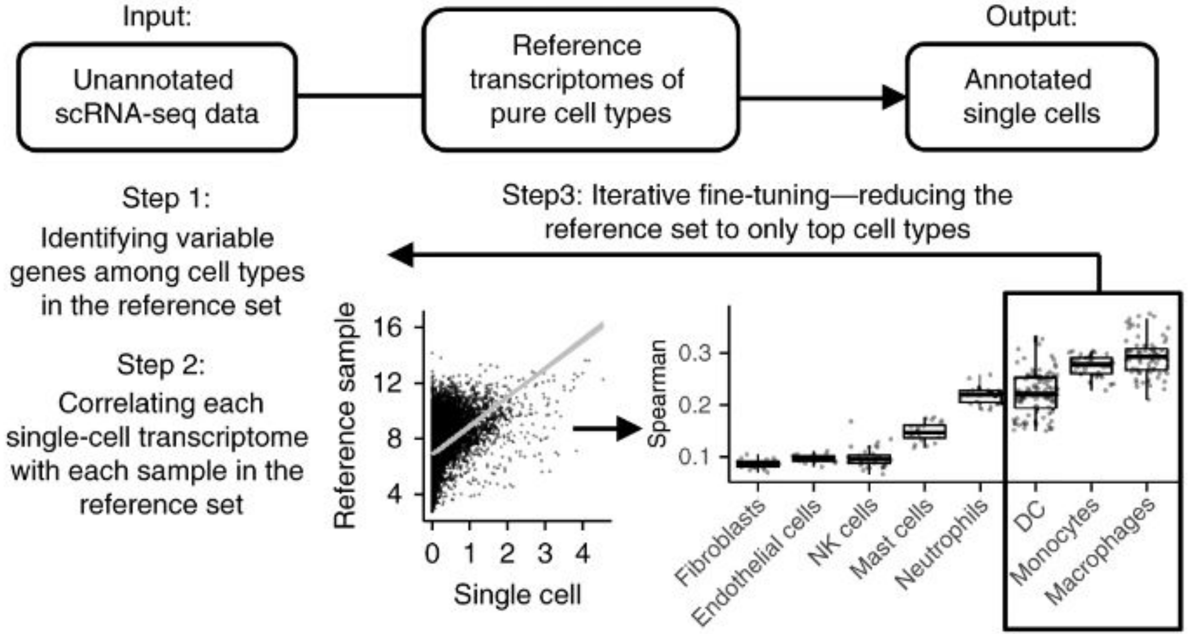


Automated cell type annotation using Bioconductor

Bioconductor packages: SingleR, celldex
Documentation: http://bioconductor.org/books/release/SingleRBook/

```
ref <- BlueprintEncodeData()

pred <- SingleR(test = sce,
     ref = ref,
     labels = ref$label.main)
```

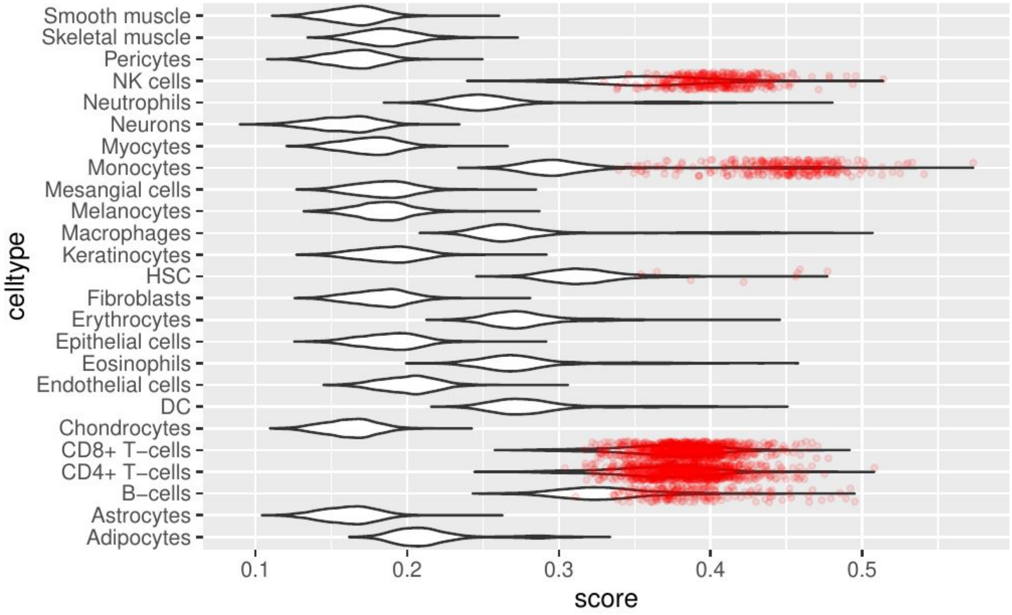


Aran et al. Nature Immunology (2019)

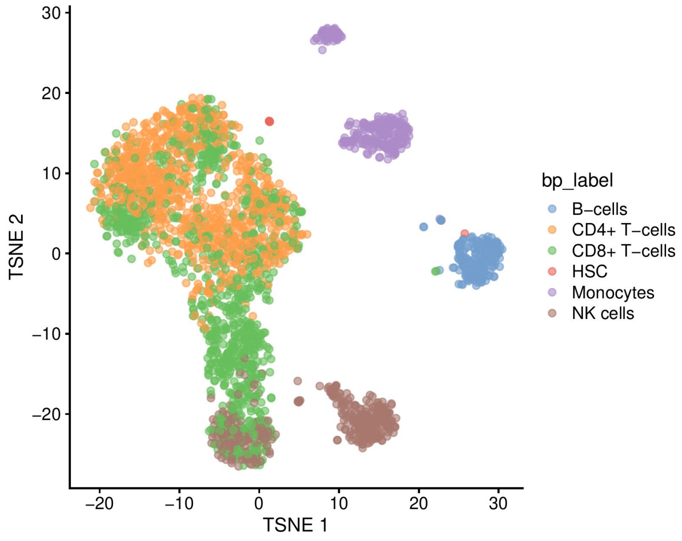
SingleR returns prediction scores and cell type labels

Scores for cell type assignment

Scores for assigned labels are indicated in red



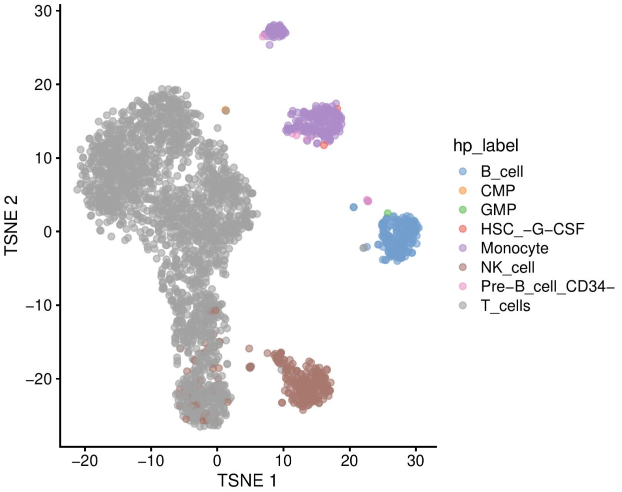
Result of automated cell type assignment



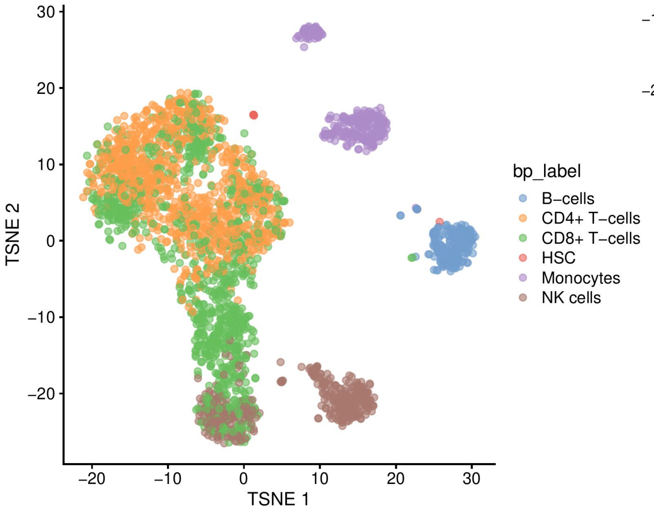
Reference: BlueprintEncodeData()

Dependent on the reference, the predictions may change

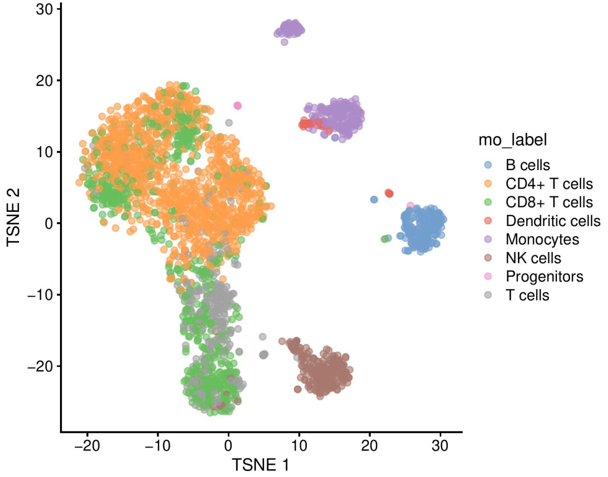
HumanPrimaryCellAtlasData()



BlueprintEncodeData()



MonacoImmuneData()



Take home messages for cell type assignment

Automated cell type assignment:

- Works well for common cell populations sequenced with whole transcriptome sequencing.
- Does not work well if you enrich for rare cell populations (NKT cells, atypical B cells)
- Does not work well for other sequencing approaches (BD Rhapsody, amplicon based).

Recommendations from my own experience:

- If you have many different cell types, split the data into subpopulations (B cells, T cells, tumor cells…). Independent subsequent analysis.
- After annotating, save an annotated intermediate object for downstream analysis.

Outline of the lecture

1. Cell types and phenotypes
2. **Interactions of immune cells**
3. Antigen specificity

Ligand - Receptor interactions

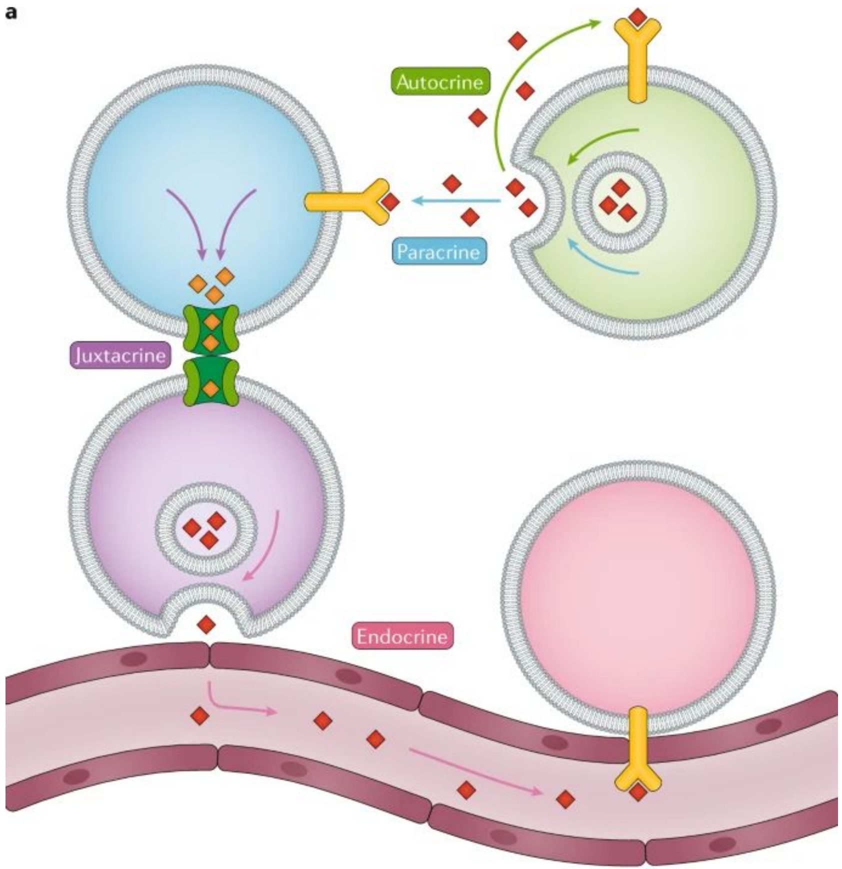
Here: Receptors and ligands encoded in the germline

Experimental measurements:

- Mass spectrometry of complexes
- Binding assays
- Affinity measurements

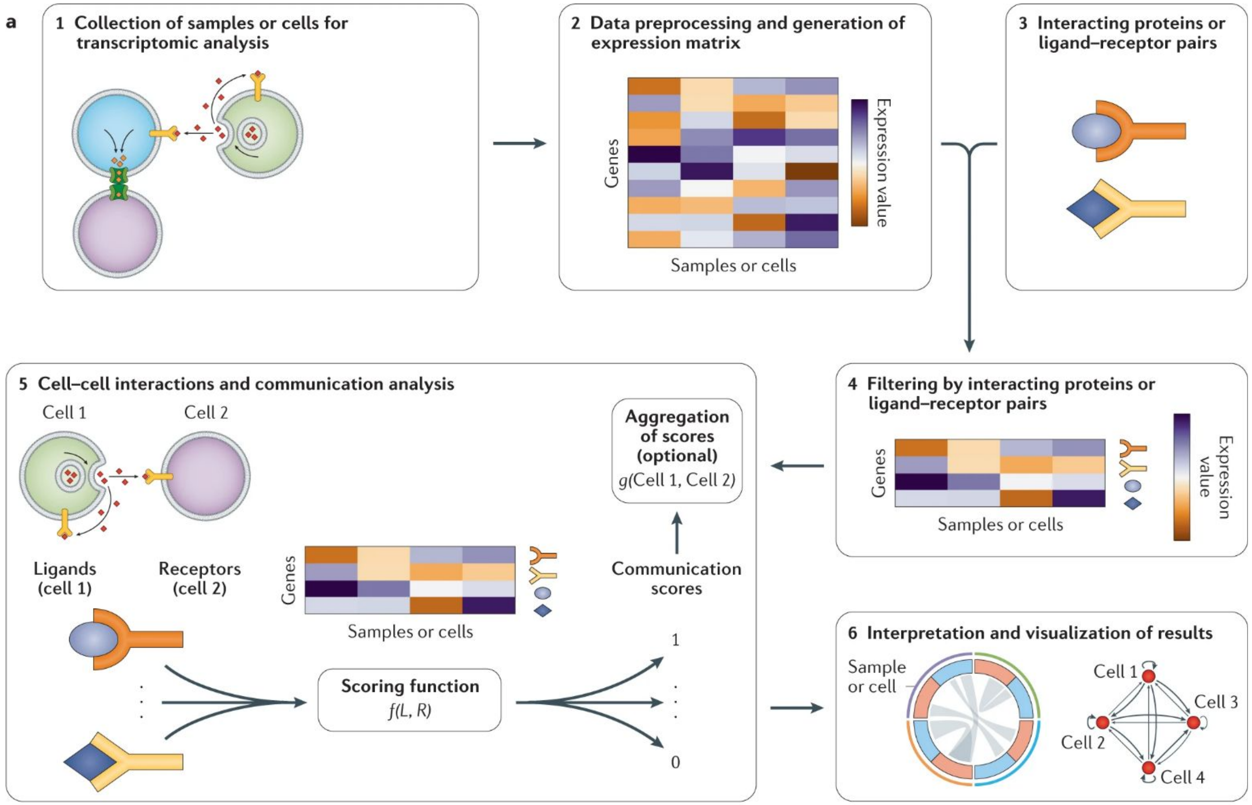
Databases of ligands and receptors:

- Cellinker
- CellChat
- CellPhoneDB
- iCELLNET
- …



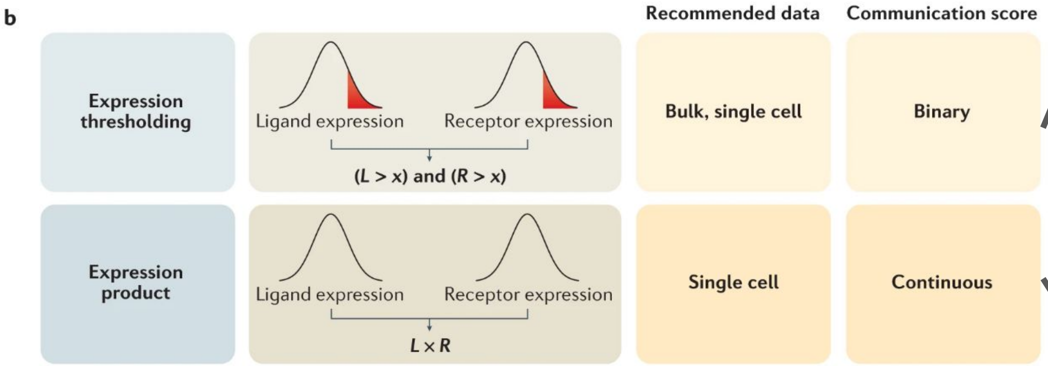
Armingol et al. Nat Rev Genetics (2021)

Workflow for scoring receptor-ligand interactions



Armingol et al. Nat Rev Genetics (2021)

Different scoring methods



	L (cell 1)	R (cell 4)	Communication scores
	1	1	1
	0	0	0
	0	1	0
	1	1	1
	1	0	0

Expression thresholding

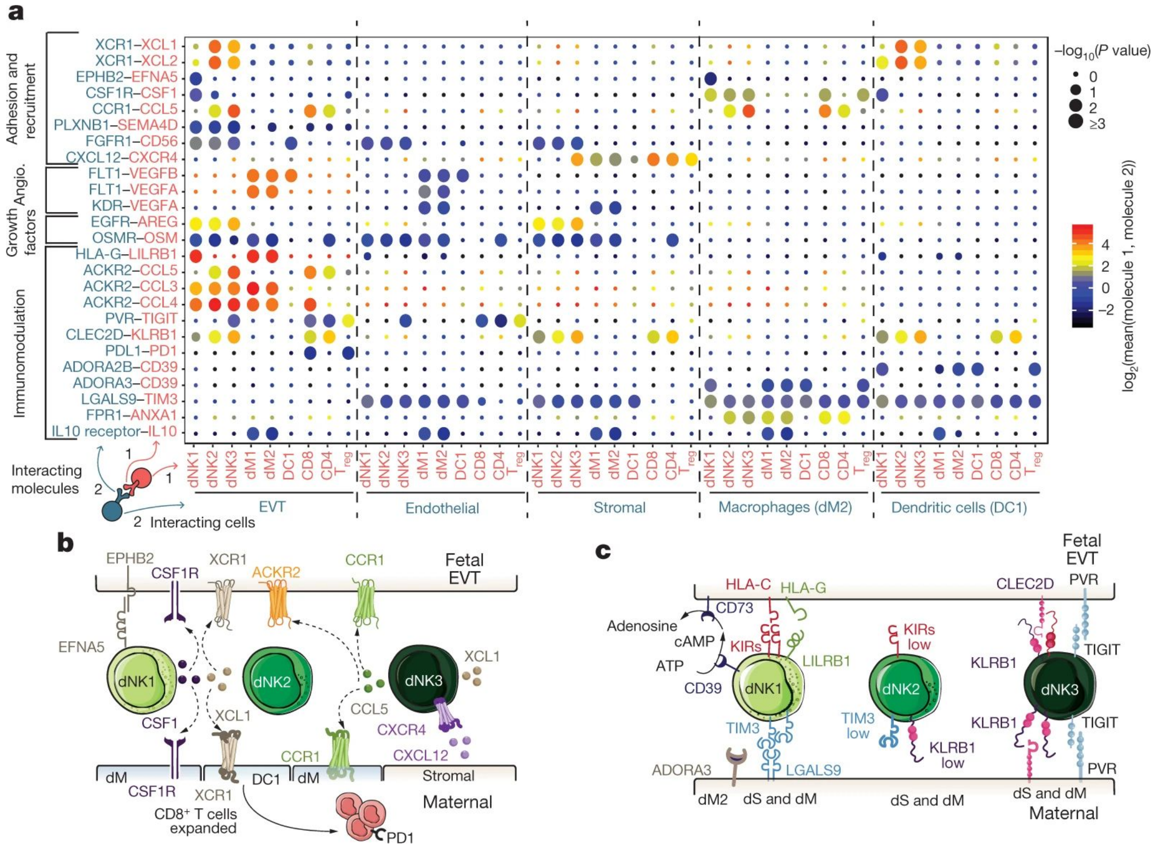
	L (cell 1)	R (cell 4)	Communication scores
	3.81	3.46	13.17
	3.17	3.17	10.05
	3.00	6.91	20.72
	4.00	6.91	27.63
	5.32	3.17	16.87

Expression product

Please see reference for complete list of tools and methods.

Armingol et al. Nat Rev Genetics (2021)

Application: Identifying receptor-ligand interactions between cell types



Vento-Tormo (2018)

Take home messages for receptor-ligand interactions

Choice of database:

Specific receptor-ligand databases for specific research communities (mouse bone marrow microenvironment, human immunology, …).

You can create your own database of receptor-ligand interactions of interest!

Common problem:

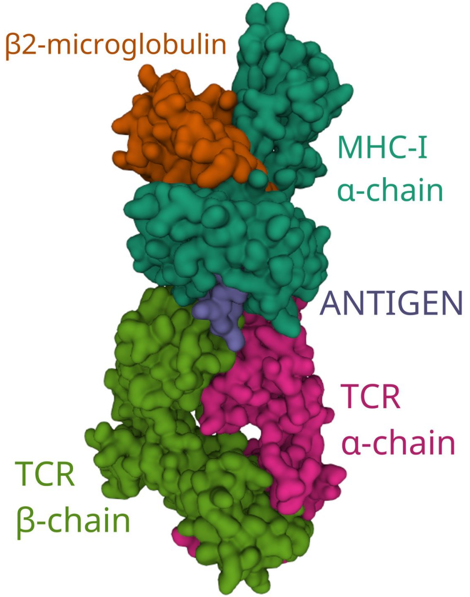
- one ligand interacts with many different receptors
- identifying differentially expressed ligands does not lead to identification of exclusive receptor-ligand pair - extensive follow up with alternative methods needed

Outline of the lecture

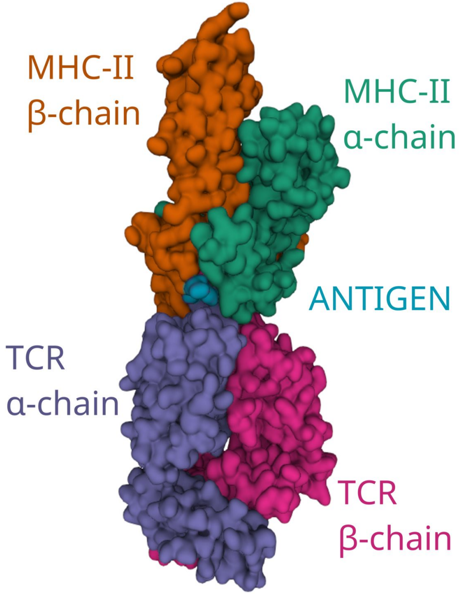
1. Cell types and phenotypes
2. Interactions of immune cells
3. **Antigen specificity**

Adaptive immune receptors are specific for a particular antigen

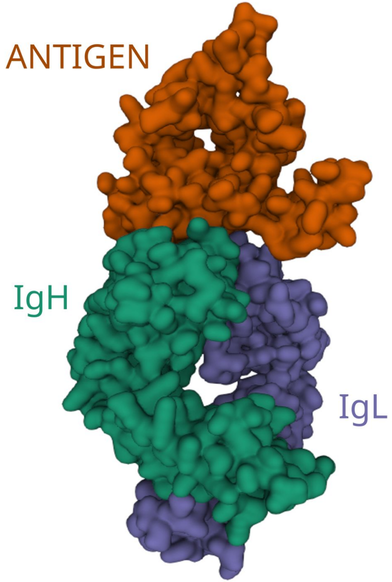
MHC-I-antigen-TCR binding



MHC-II-antigen-TCR binding



Antibody-antigen binding



Adaptive immune receptor repertoires (AIRR)
V(D)J recombination → Sequence diversity of antigen receptors

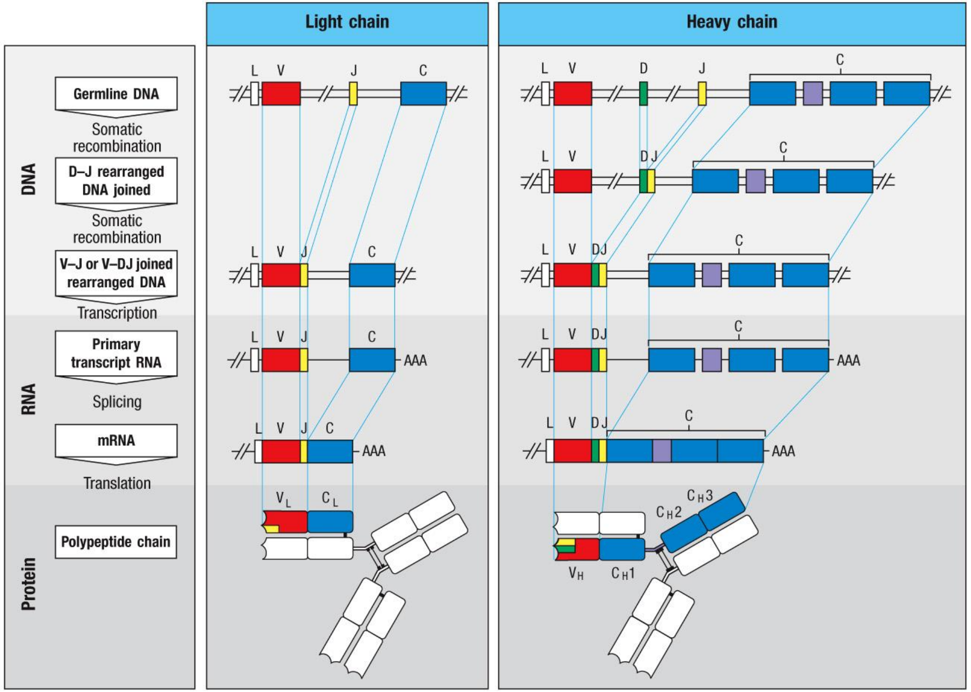


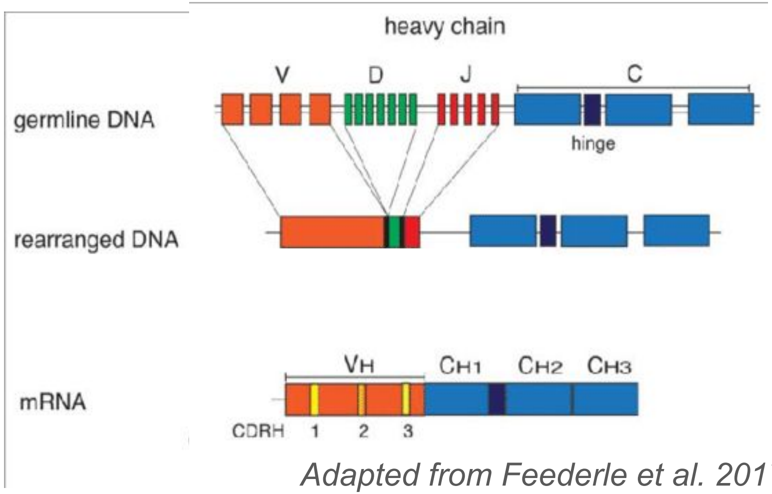
Figure 5.3 Janeway's Immunobiology, 9th ed. (© Garland Science 2017)

Number of functional gene segments in human immunoglobulin loci

Segment	Light chains		Heavy chain
	κ	λ	H
Variable (V)	34–38	29–33	38–46
Diversity (D)	0	0	23
Joining (J)	5	4–5	6
Constant (C)	1	4–5	9

Figure 5.4 Janeway's Immunobiology, 9th ed. (© Garland Science 2017)

Adaptive immune receptor repertoires (AIRR)
V(D)J recombination → Sequence diversity of antigen receptors

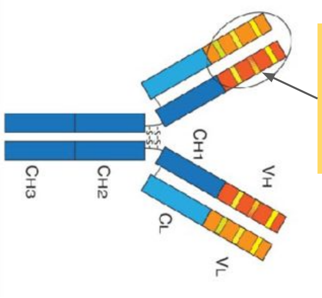


Adapted from Feederle et al. 2017

Number of human immunoglobulin genes (heavy chain)

V genes	ca. 45
D genes	23
J genes	6

Antibody

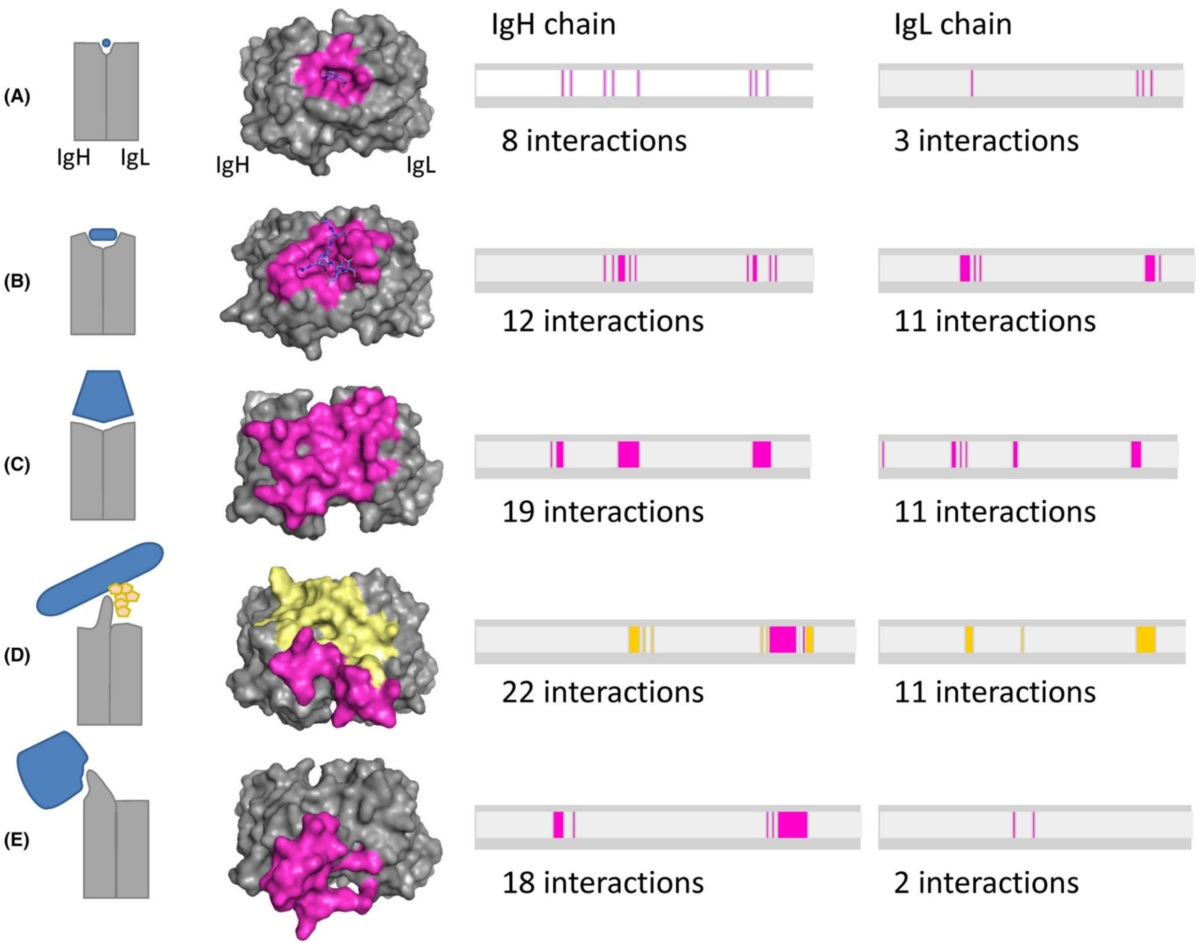


CDR - complementarity determining regions

→ **combinatorial diversity!**

Database: www.imgt.org (Lefranc et al. 2001)

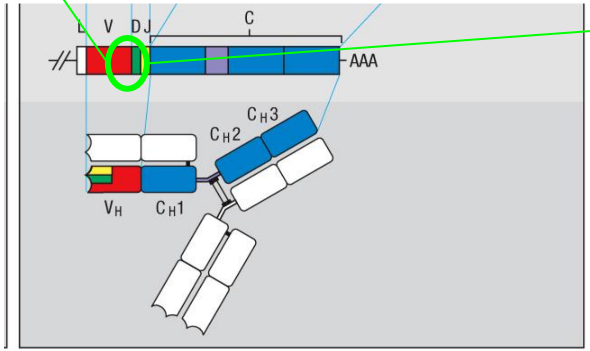
Diversity of antibody-antigen interactions



Imkeller, Wardemann (2018)

Annotating immunoglobulin sequences

```
                                            ----------------><------------------------------CDR3-IMGT----------------------------->
                                             T  A  V  Y  F  C  A  R  D  L  S  C  T  S  T  T  T  C  H  R  P  L  K  T  N  Y  G  M  D  V
                          Query_1     362   ACGGCTGTTTATTTCTGTGCGAGAGATTTGAGTTGTACTAGTACTACCACCTGCCATAGGCCGTTGAAGACAAACTACGGTATGGACGTC  451
V  94.9% (280/295)        IGHV1-3*04  271   ........G....A..........------------------------------------------------------------------  295
                                             T  A  V  Y  Y  C  A  R
D  85.0% (17/20)          IGHD2-2*02  10    --------------------------------------.....G.....G....T...--------------------------------  29
J  100.0% (50/50)         IGHJ6*02    13    -------------------------------------------------------------------------.................  29
```

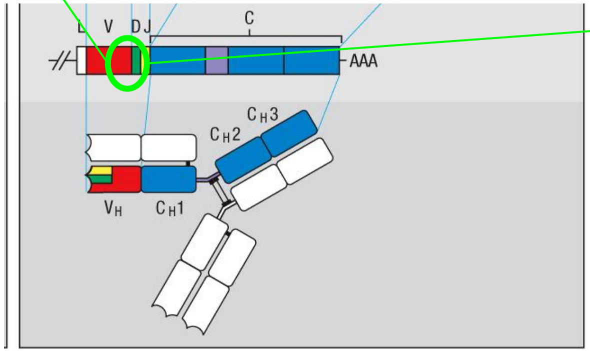


Features:

- V,D,J usage
- CDR/FWR

Annotating immunoglobulin sequences

```
                                                    ----------------><--------------------------------CDR3-IMGT------------------------------->
                                                     T  A  V  Y  F  C  A  R  D  L  S  C  T  S  T  T  T  C  H  R  P  L  K  T  N  Y  G  M  D  V
                           Query_1      362   ACGGCTGTTTATTTCTGTGCGAGAGATTTGAGTTGTACTAGTACTACCACCTGCCATAGGCCGTTGAAGACAAACTACGGTATGGACGTC  451
V  94.9% (280/295)   IGHV1-3*04   271   ........G....A..........-----------------------------------------------------------------  295
                                                     T  A  V  Y  Y  C  A  R
D  85.0% (17/20)     IGHD2-2*02   10    --------------------------------------.....G.....G....T...--------------------------------  29
J  100.0% (50/50)    IGHJ6*02     13    -------------------------------------------------------------------------.................  29
```



Features:

- V,D,J usage
- CDR/FWR
- Somatic hypermutations

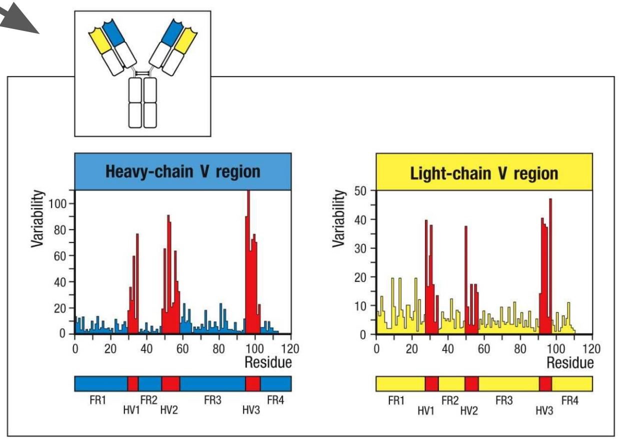


Figure 4.6 Janeway's Immunobiology, 9th ed. (© Garland Science 2017)

AIRR annotation tools:

IgBLAST (Ye et al. 2013), Immcantation (Kleinstein, Vander Heiden, Gupta, Yaari et al.), …

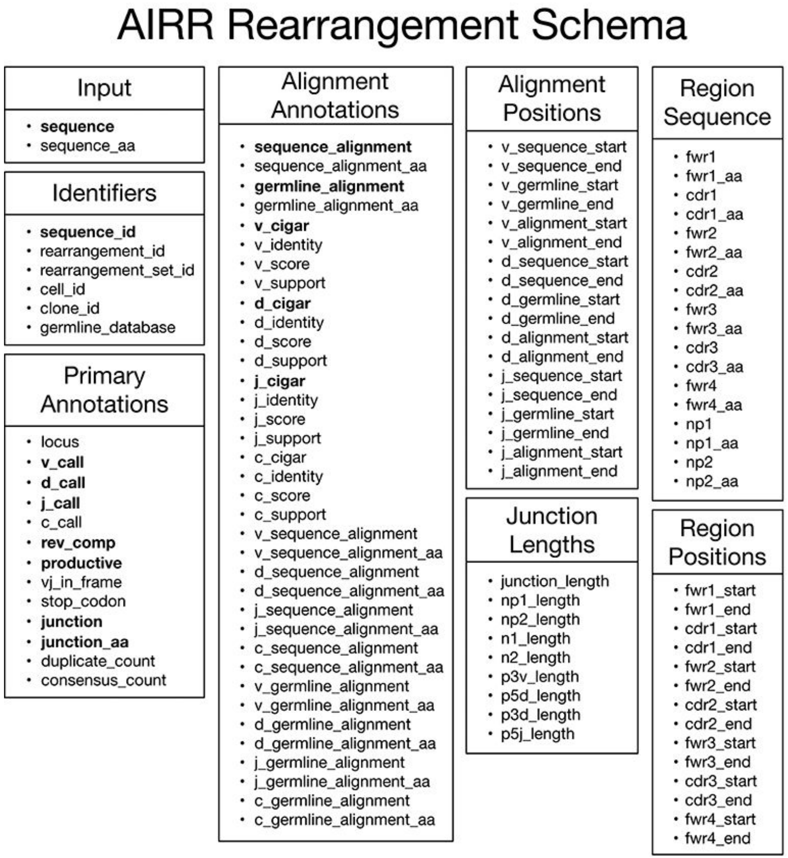
AIRR data format

Standard format for annotating adaptive immune receptor sequences

https://docs.airr-community.org/en/stable/api/adc.html#datacommons

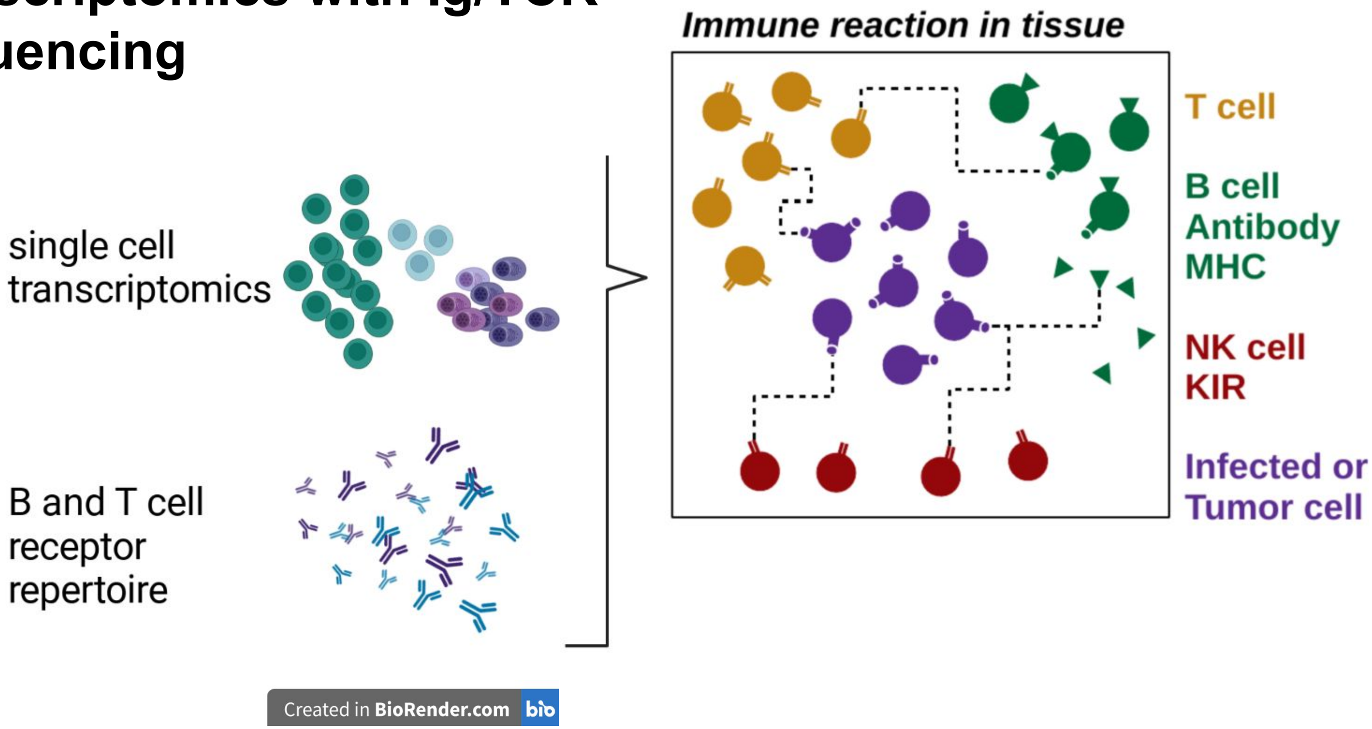
Annotation by Cellranger in AIRR format:

`airr_rearrangement.tsv`

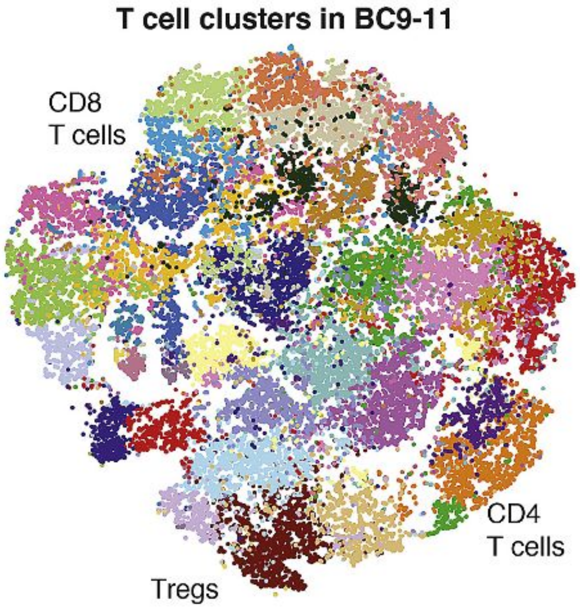


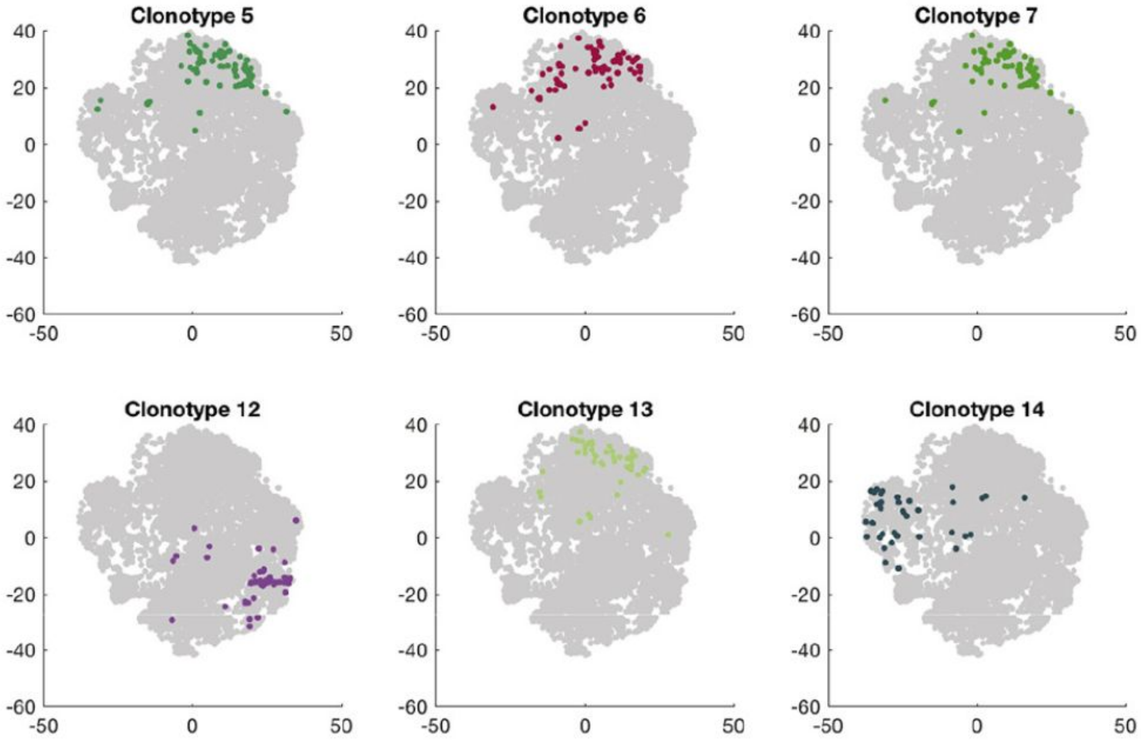
Vander Heiden et al. (2018)

Integrating single-cell transcriptomics with Ig/TCR sequencing



TCR Clonotypes and single cell transcriptomic data





Azizi et al. Cell (2018)

Take home messages for AIRR and antigen specificity

The receptor repertoire is VERY diverse (10^{12}):

- when sequencing AIRR from the blood, isolate antigen-specific or activated cells
- use the common AIRR data format for best compatibility with existing analysis tools

Prediction of MHC-bound peptides:

- works well for well-studied alleles and MHC class I
- reliable methods for human MHC class II prediction still under development

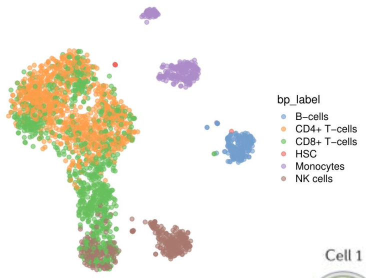
Summary

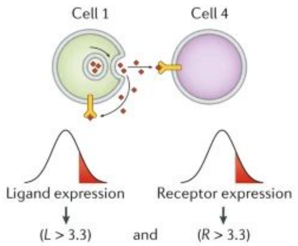
Immune cell types and interactions

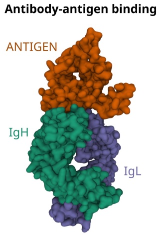
- automatic annotation immune cell types (or not)
- how to score receptor ligand-receptor interactions in gene expression datasets

Antigen specificity and adaptive immune receptor repertoires

- annotation of adaptive immune receptor repertoires
- AIRR features and clonotypes







Further reading

Cell type annotation in OSCA: https://bioconductor.org/books/release/OSCA/

SingleR book: http://bioconductor.org/books/release/SingleRBook/

Review cellular interactions:

Armingol, E., Officer, A., Harismendy, O. *et al.* Deciphering cell–cell interactions and communication from gene expression. *Nat Rev Genet* 22, 71–88 (2021). https://doi.org/10.1038/s41576-020-00292-x

Review B/T cell repertoires:

Philip Bradley and Paul G. Thomas. Using T Cell Receptor Repertoires to Understand the Principles of Adaptive Immune Recognition. Annual Reviews Immunology (2019). https://doi.org/10.1146/annurev-immunol-042718-041757

Katharina Imkeller, Hedda Wardemann. Assessing human B cell repertoire diversity and convergence. Immunological Reviews (2018). https://doi.org/10.1111/imr.12670