**Report on choosen CAR-T targets**

EGFR, CD7, MSLN, and CD19 have been evaluated as potential CAR-T targets.

For each target, the properties that have been evaluated are:

* Isoforms expression in cancer and healthy tissues
* Protein topology
* Cellular localization
* Epitope colocalization, when possible (missing)

The script (Rmarkdown) generating the results is in: cartcontent/scr/ 02\_Targets\_evaluation\_figures\_paper\_2024.Rmd.

**Isoforms expression in cancer and healthy tissues**

EGFR 🡪 higher in cancer, but not absent in healthy tissues.

CD7 🡪 it is normally expressed on T cells and NK cells (source [the protein atlas](https://www.proteinatlas.org/ENSG00000173762-CD7/single+cell+type)). However, in TCGA the only blood related cancers are acute myeloid leukemia and diffuse large B cell lymphoma. So we cannot verify the expression of CD7 in T cell related malignancies.

A screenshot of a computer

Description automatically generated

Based on the GTEX data, CD7 is also expressed in lung, small intestine, spleen, stomach. Most likely, the reason why it is expressed in such tissues is that there are infiltrating T cells in the tissue. We could check the literature, to see if there is agreement with the fact that CD7 is only expressed on T cells or NK cells. In addition, we could check how the expression of CD7 correlates with CD3 or CD56 espression (T cell and NK markers)?

MSLN 🡪 in both cancer and healthy tissues.

CD19 🡪 it is used as target for both hematological and solid malignancies.

It is expressed on healthy small intestine, spleen, stomach, testis.

Same as for CD7: the reason why it’s expressed on non-haematological tissues is most likely because there are B cells infiltrating in such tissues. Also in this case, we should check if literature is in agreement with CD19 being only expressed by B cells. We can also check how the expression of CD19 correlates with CD20 or other known B cell markers.

Conclusion: it is hard to find a target that is not expressed in healthy tissues. Maybe we can focus on bringing CD19 or CD7 as example?

**Protein topology & cellular localization**

Protein topology is line with cellular localization: for CD7, CD19, and EGFR, variants with a TM domain are also predicted to be on the cell membrane by DeepLoc2. Regarding MSLN, none of the variants has a TM domain. DeepLoc2 still predicts them as membrane bound. This can be explained by the fact that MSLN is anchored to the cell membrane by glycosylphosphatidylinositol (GPI) anchors, and therefore does not need a TM domain. It would be interesting to predict possible GPI anchors (tool [NetGPI](https://services.healthtech.dtu.dk/services/NetGPI-1.1/)).

Example: EGFR (the canonical isoform is highlighted in the plots below).

Both ENST00000275493 and ENST00000455089 have a TM domain. They are also predicted to be localized on the cell membrane by DeepLoc2. Other variants do not have the TM domain, and are predicted to be extracellular.

A colorful squares with white text

Description automatically generated



A graph with different colored bars

Description automatically generated with medium confidence

