# Chapter 6 - Conclusions and Future Directions

## 6.1 Conclusions

In this thesis, I presented results from our single cell-resolution characterization of the human maternal-fetal interface during early pregnancy. Single-cell transcriptomes derived from droplet-based and plate-based scRNA-seq of maternal decidua and fetal placenta were analyzed in conjunction with whole genome sequencing of maternal and fetal genomic DNA, bulk RNA-sequencing of decidual tissues, FACS index sorting and CyTOF datasets, established macrophage gene signatures and public gene expression datasets, and curated gene sets associated with complications of pregnancy and fertility.

I first demonstrated that while determination of the maternal or fetal genetic identities of decidual and placental cells requires further optimization before it can be achieved through scRNA-seq-based variant detection alone, we can successfully perform maternal/fetal assignments for these cells in cases where maternal and fetal genotypes are known, using the recently developed demuxlet algorithm. This unbiased determination of maternal/fetal origin enabled us to independently evaluate the relative presence of maternal and fetal cells in the placenta and decidua. At the same time, it also improved our annotation of cell types at the maternal-fetal interface by clarifying the genetic origins of specific cell populations of interest, including mononuclear phagocytes localized to the placenta found to be of maternal origin, and an unknown cell population found in both decidua and placenta determined to be of fetal origin.

I then showed that analyzing scRNA-seq data in conjunction with matched bulk RNA-seq data generated on the same samples represents a potential strategy by which to identify genes susceptible to single-cell dissociation protocol-induced perturbations in expression. We found that *HSPA1B*, a gene highly upregulated in a decidual T cell subpopulation identified by plate-based scRNA-seq, was not significantly expressed in bulk RNA-seq of the same decidual tissues. This suggests that *HSPA1B* and the T cell subpopulation characterized by its expression are potential scRNA-seq-associated artifacts, and we therefore excluded these cells from subsequent functional annotation and analysis.

Next, I showed how superimposition of gating identities from FACS index sorting onto scRNA-seq data enabled us to validate the transcriptome-based annotations of decidual immune cells sequenced by the plate-based SmartSeq2 protocol. In particular, FACS gating enabled us to confirm the existence of two distinct decidual mononuclear phagocyte subsets distinguished by surface expression of HLADR and

CD14 and by high and low expression of the *CD11C* gene. Using high-dimensional CyTOF, we were able to further distinguish these myeloid cell populations based on CD11C protein expression, as well as independently validate and further describe the novel dNK and distinct T cell populations that were previously annotated in the scRNA-seq data.

To further characterize the maternal mononuclear phagocyte populations at the maternal-fetal interface, we performed GO and Reactome term enrichment analysis on the upregulated genes in this population and paired single-cell transcriptomic data from these subsets with M1/M2 macrophage signatures and placenta- and endometrium-enriched genes identified from datasets deposited in the Human Protein Atlas. These analyses enabled us to determine that the mononuclear phagocytes collectively upregulate genes involved in tissue remodeling, modulation of local inflammatory responses, regulation of ROS production and detoxification, and lipid sensing and metabolism. Overall, we show that maternal decidual and placental mononuclear phagocytes do not exhibit classical M1/M2 macrophage polarization and, along with several other decidual resident immune subsets, are characterized by their elevated expression of a number of genes enriched in the placenta and endometrium, suggesting their adaptation of tissue-specific functions that heavily tailor them to the unique demands and microenvironment of the maternal-fetal interface.

Finally, the curation of genes associated with complications of pregnancy and conditions influencing fertility and the intersection of these gene sets with our decidual and placental scRNA-seq data enabled us to better understand how specific cell types at the maternal-fetal interface potentially contribute to driving these diseases or conditions. In particular, we show that fetal trophoblasts express a large proportion of the fetal genes previously associated with preeclampsia, while maternal smooth muscle and endothelial cells preferentially express a number of the maternal preeclampsia-linked genes, reinforcing the complementary roles and potential interplay between different cell populations in regulating trophoblast cell differentiation and invasion and maternal vascular adaptation, and in turn, their respective contributions to the complex pathogenesis of preeclampsia and similar disorders associated with defective placentation. Meanwhile, we show that among the genes influencing fetal growth and birth weight, maternal expression of IGFs and their binding proteins is concentrated among the decidual stromal cell subsets, while fetal expression of IGF axis genes is distributed among trophoblast cell, fibroblast, and Hofbauer cell populations, suggesting that IGF regulation of fetal growth is potentially mediated in an autocrine and paracrine fashion between the decidua and placenta and

involves similarly complex mechanisms of cross-regulation and coordination among a number of cell types at the maternal-fetal interface.

Taken as a whole, this series of analyses integrating single-cell transcriptomics with additional experimental approaches and datasets have enabled us to validate or to uncover novel findings and interesting biological aspects of the maternal-fetal interface that may not have been achievable through analyses of scRNA-seq data in isolation.

## 6.2 Future directions

### *6.2.1 Further improvements and applications of single-cell genotyping*

While we were able to successfully infer the maternal or fetal identities of single cells from droplet-based scRNA-seq data through the use of demuxlet and maternal and fetal WGS, it would ultimately be desirable to develop a method to robustly genotype single cells via scRNA-seq read data alone. This would eliminate the need to generate additional genotype data from all genetically distinct individuals associated with a given sample, a practical benefit that would scale particularly as the number and variety of tissues and samples to be explored by scRNA-seq increases. There are many conceivable scenarios in which the number or identity of distinct individuals potentially contributing cells to a sample is unknown *a priori*, making it impossible to obtain the genomic DNA or genotypes needed for single-cell genotyping by demuxlet. Samples involving organisms lacking well-annotated DNA reference genomes, or more generally, any sequencing experiment in which there are potential sources of contaminating cells, present particular challenges. A method which would be able to robustly infer genetic distinctness among all cells in a sample in an unbiased manner would likely significantly advance efforts to dissect such systems.

Meanwhile, single-cell genotyping could be leveraged to further investigate certain questions of specific relevance to the maternal-fetal interface. Genomic imprinting has been observed in several genes expressed in the placenta and found to play various roles in regulating processes at the maternal-fetal interface[213]. In fact, placenta-derived *IGF2*, linked to fetal growth, has been shown to be maternally imprinted (with only the paternally inherited copy expressed)[214], and similarly, other genes involved in fetal growth and enriched in the placenta, including *MEST, MEG3,* and *PHLDA2*, have also shown to be expressed in an imprinted manner. An endeavor to identify all imprinted genes at the maternal-fetal interface would prove useful in understanding the functional and phenotypic consequences of imprinting

at specific genes, and perhaps more broadly, the evolutionary basis for imprinting to persist as a gene regulatory mechanism at the maternal-fetal interface. Single-cell genotyping could be adapted to comprehensively identify imprinted genes by determining which genes expressed in decidual and placental cells exhibit allelic imbalance, and furthermore, help clarify whether the imprinting of a gene of interest occurs exclusively in specific cell types or cell states or during specific stages of pregnancy. Single-cell genotyping could also be applied to studying microchimerism, another interesting extension of maternal-fetal biology. In cases of fetal microchimerism, cells of fetal origin pass through the placenta to the maternal circulation and establish lineages in maternal tissues, while in maternal microchimerism, maternally derived cells are passed to the child and shown to persist into adulthood[215]. Although constituting a small minority of cells in their host (e.g. in maternal T cells, up to 2.7 per 100,000 cells[216]), microchimeric cells have been observed in the immune cell populations of 39% of women after birth[216], as well as in women with history of induced or spontaneous abortion[217]. Both maternal and fetal microchimerism have been observed in hematopoietic cell lineages, as well as in cardiac myocytes, hepatocytes, and differentiated cells in other tissues[218], and have been variously associated with autoimmune conditions such as rheumatoid arthritis[219], neonatal lupus[220], complications of pregnancy[218], and cancer[221]. The ability to incorporate unbiased single-cell genotyping into the study of additional tissue samples, which are already being actively studied at single-cell resolution through initiatives such as the Human Cell Atlas (HCA) and Developmental Cell Atlas, may provide new insights into determining the extent of microchimerism events and investigating their potential functional impacts on disease.

### *6.2.2 Further characterization of identified cell types at maternal-fetal interface*

In this study, a number of cell populations, including novel cell types or states, were described in the decidua and placenta, with identification of their distinct gene and surface protein markers, functions, and relevance to disease. Critical to robustly validating these cell populations and their proposed functions would be to further observe their predicted gene and protein expression patterns and cell-cell interactions *in vitro* or *in vivo*. Anatomical and cellular differences in animal models of pregnancy, in combination with the limitations of many *in vitro* models, remain challenges to be addressed. However, the recent development of endometrial organoids, which are capable of functionally recapitulating early pregnancy[222], may prove useful for further functional studies of cells in the decidua, such as the distinct

stromal cell populations and their respective secretory characteristics identified in this study. Meanwhile, intact decidual and placental samples can be further analyzed using RNAScope[223], a novel *in situ* hybridization assay which facilitates the examination of selected RNA biomarkers within cells in fresh frozen or formalin-fixed, paraffin-embedded tissue sections at single-cell resolution. RNAScope represents a valuable means of validating gene signatures and predicted cell-cell interactions in cell populations, while also preserving the spatial architecture of the tissue being studied, thus enabling it to control for dissociation protocol-associated perturbations in gene expression while also providing a new layer of information about the spatial organization of the cells of interest, which will be relevant to further contextualizing their functions and proposed interactions.

At the same time, additional analyses at the single-cell transcriptome level will be useful in further annotating the cell populations featured in this study. For example, in order to define relationships between the different identified trophoblast subsets, pseudotime algorithms such as Monocle[224] were employed to infer a developmental trajectory among the trophoblast cells, defining their differentiation from progenitor cells and cytotrophoblast cells into the SCT and EVT that facilitate placentation, and in turn, identifying the transcriptional changes that potentially underlie this process which were previously not well understood (Vento-Tormo R, Efremova E, et al., under submission). In this thesis, we further described the three maternal mononuclear phagocyte subsets identified in the decidua and placenta, determining that dMP1 and plMP exhibit a monocyte-derived origin and that dMP2 exhibits a markedly tissue-resident phenotype. We could employ similar approaches for trajectory inference to dissect the nature of the monocyte to tissue-resident macrophage transition that characterizes these subsets and thus better define their potential interrelationships.

Finally, extending our single-cell analyses from early pregnancy towards the study of nonpregnant endometrium, as well as near-term or term decidua and placenta, would enable the construction of a comprehensive atlas of the cell states, biology, and immunology of the maternal-fetal interface over the various time points of pregnancy, and would also enable comparative analysis and inference of how these states change over time. In particular, a longstanding question in the characterization of maternal immune cell subsets in pregnancy is whether these cells reside in the endometrium or decidua prior to pregnancy, or alternatively, at what stages before or during early pregnancy they are recruited to the maternal-fetal interface[2]. Single cell-level mapping of various states of the nonpregnant endometrium (proliferative and secretory phases) in comparison with the early pregnancy decidua could prove pivotal

in addressing this gap in our knowledge. More broadly, a single cell-resolution understanding of the successive stages of pregnancy would undoubtedly facilitate the identification of novel cell states and cell functions in addition to the refinement of known ones.