Ever since their first introduction, medicines have been a fundamental component of daily life. Thousands of drugs have already been discovered and significantly improved patient outcomes. However, one of the most difficult challenges associated with pharmacotherapy is the interpatient variability in drug response; both the effectiveness of drugs and their safety profile differ markedly across the population. This has led to the emergence of novel approaches in pharmaceutical healthcare. Two notable examples are the shift towards more personalized medicine and more targeted therapies. These novel approaches have both been advanced by leveraging genomic insights from sequencing data. The unique genetic profile of a patient can assist in the prescription of tailormade treatments for a specific individual, with the promise of more effective, efficient and patient-centered healthcare. This is the subject of research in the field of PGx, representing a cornerstone of personalized medicine. Additionally, leveraging antibody DNA sequences from convalescent or immunized individuals can also facilitate the development of fully human, targeted mAb therapeutics, playing a transformative role in mAb drug development. However, despite the promise of these approaches, current methods in both PGx and therapeutic mAb development face limitations and could benefit from further optimization. Therefore, this dissertation set forth two primary research aims to address these areas of opportunity. The first aim of this dissertation was to investigate several alternative approaches for obtaining more comprehensive and phased data to improve PGx genotyping. The second aim was to explore a novel workflow to tackle current problems with single B cell screening and the generation of antibody sequences in the context of mAb development.

Part A of this dissertation addressed the first aim. During the last decades, the importance of personalized medicine has become increasingly apparent. Directly linked to that is the need for accurate genotyping assays to determine the PGx profile of patients. While current genotyping assays perform quite robustly, they struggle with complex genes and fall short in identifying unknown variants and defining full haplotypes. Many clinically relevant pharmacogenes, such as *CYP2D6*, are accompanied by pseudogenes with sequence homology, posing severe challenges for genotyping. Traditional methods risk capturing sequences from the pseudogenes in addition to the functional genes, leading to mapping errors, low variant detection rates and a high number of false positives. Therefore, sequencing methods that provide the sequence of an entire gene could be useful to overcome this problem. Furthermore, as a gene's function is determined by the combination of all variants per allele, the importance of phased, full-gene sequencing for accurate phenotype predictions cannot be ignored. Therefore, **Chapters 3** and **4** of this dissertation assessed alternative genotyping assays aimed at obtaining

accurate, phased, full-gene sequences for several pharmacogenes, including the complex *CYP2D6* gene, by using targeted linked- and long-read sequencing technologies.

Chapter 3 aimed to generate complete allele-specific sequences of *CYP2D6* by optimizing a targeted amplification-free LRS method and developing an improved analysis pipeline. Target enrichment was obtained by applying the nCATS strategy, or nCATS combined with AS on 5 µg of DNA from three well-defined cell lines. LRS was performed using an ONT GridION device, followed by data analysis using the developed CoLoRGen pipeline, aiming to simultaneously detect both large structural variants as well as small variants. The use of the AS software in addition to the nCATS enrichment did not consistently result in a higher on-target depth, limiting its advantage in this context. The nCATS strategy on itself did not reach the anticipated 100X depth either, which was probably due to the occurrence of sequenceable background DNA as well as the presence of unsequenceable non-adapter-ligated DNA molecules on the flow cell. This low target depth comprises one of the main disadvantages of the nCATS enrichment method in the PGx context. However, although the nCATS enrichment proved suboptimal, the nCATS-CoLoRGen assay demonstrated accurate genotyping for the complex *CYP2D6* locus. When the minimum depth of 16X for each allele and three reads covering the breakpoints of structural variants were obtained, correct star-alleles could be assigned to *CYP2D6* and the *CYP2D6-CYP2D7* hybrid for the three cell lines. Moreover, by generating a complete consensus sequence of the genes, the CoLoRGen pipeline provided direct evidence of the presence of *CYP2D6-CYP2D7* large structural variants and smaller SNVs and INDELs that go undetected by other current methods. By enabling the possibility to detect and phase *de novo* mutations in addition to known large structural and small variants, this genotyping assay allows for more accurate gene function predictions, which should be likewise extendable to other genomic regions of interest. However, the practical application of this assay is hindered by the inherent limitations of the selected enrichment strategies, resulting in low on-target sequencing depths and necessitating the use of one costly flow cell per patient.

Chapter 4 assessed the potential of a linked-read enrichment approach, called TLA, in terms of accurate phasing. For this purpose, four pharmacogenes, *CYP2D6*, *CYP2C19*, *CYP1A2*, and *BRCA1*, were enriched with different generic primer pairs and sequenced both with Illumina and Nanopore sequencing technologies. The four most important conclusions can be summarized as follows: i) due to drawbacks inherent to TLA, three out of the four genes were not entirely covered with either sequencing method; ii) this resulted in incomplete genotypes for these genes, with only a minor part of the variants that could be called and phased correctly compared to reference sets; iii) overall, variant calling was more accurate in the Illumina datasets, whereas more accurate phasing was obtained with Nanopore sequencing; and iv) only for the complex *CYP2D6* gene, the long Nanopore reads

improved the variant calling compared with Illumina sequencing, generating the only correct haplotype. Altogether, due to its inherent drawbacks, the TLA workflow is considered unsuitable as a generalized assay for reliable genotyping and phasing of important pharmacogenes. Although patient-specific primer design based on prior knowledge of heterozygous SNV positions could improve these results, this seriously complicates the process and increases the required cost and time-commitments, rendering the application of TLA as a PGx assay impractical.

The findings in **Chapters 3** and **4** of this dissertation suggest that, although possibly useful in other contexts, the examined technologies are not sufficiently refined for accurate, cost-effective PGx testing. Nevertheless, given the sustained value of LRS technologies for their ability to provide comprehensive and phased PGx genotyping, targeted LRS sequencing remains a promising method to advance the field of PGx. Especially since Nanopore sequencing is still a rather new method, continuously improving and decreasing in cost, this technology still holds significant potential for cost-effective PGx genotyping in the future, possibly leading to more accurate gene function predictions.

Part B focused on the second research goal by exploring a novel workflow for rapid and high-throughput screening and antibody sequence analysis of ASCs at the single-cell level in the context of mAb development. Over the years, mAbs have become increasingly apparent for a wide range of applications. Their role as therapeutic agents is growing exponentially due to their highly targeted nature, thereby limiting the occurrence of side effects. However, commonly used technologies for mAb development often contain time-consuming screening processes, which impair the rapid development of new mAb therapeutics in urgent situations such as the COVID-19 pandemic. The newer single B cell technologies have seen increased use in mAb research lately, as they enable more rapid and high-throughput screening of the antigen-specificity of human B cells to eventually generate fully human mAbs using their antibody sequence. **Chapter 6** aimed to verify if an innovative method can overcome some of the drawbacks of available screening methods and prove to be an equivalent or even better alternative for screening and analyzing single B cells. The strength of the method lies in the design of a continuous workflow for the steps of single-cell confinement, antibody secretion screening, visual inspection and retrieval of B cells of interest, and identification of their antibody sequence. By additionally leveraging small well volumes and combining 6,400 wells into a single microwell chip, this method offers a high-throughput and fast system for the screening and analysis of single human B cells. As patient-derived B cells are notably short-lived in *ex vivo* culture conditions and challenging to interrogate, the small well volumes and swift workflow offer significant advantages in this context. Moreover, culture conditions and RT-PCR protocols were additionally optimized to successfully obtain antibody DNA sequences from the single B cells of interest. This resulted in the

identification of complementary V_H and V_L DNA pairs from 43% of the retrieved single ASCs of interest, comparable with attainable efficiencies in existing protocols. Overall, the whole workflow, from blood sampling to detection of antibody sequences, could be successfully performed in approximately 1 day.

Furthermore, the potential of the novel platform for screening hybridomas was additionally evaluated, considering hybridoma technology remains widely used for mAb development but suffers from lengthy screening steps. The workflow proved capable of seeding, screening, and isolating hybridomas on a single-cell level in a high-throughput manner. Condensing these processes to approximately 1 day, this method shows a marked improvement over the 7–10 days required per screen for the traditional approach involving limiting serial dilution and repeated screening by ELISA. These results underscore the promising opportunity and advantages of this workflow for single-cell antibody screening and analysis purposes. Additionally, the workflow has broad translational potential for other cell types or even other immunological assays. Nevertheless, this technology is still in its experimental phase and requires further refinement and validation for practical applications. With further optimization, such high-throughput single-cell technologies could revolutionize mAb development in the future, enabling the discovery of highly effective therapeutic antibodies at unprecedented speeds.

To conclude, this dissertation contributes to two rapidly evolving domains of pharmaceutical healthcare. By demonstrating the promising opportunity of alternative methods in the fields of PGx and mAb development, these studies will be the starting point for exciting new research in both fields. In the field of PGx, more comprehensive targeting sequencing technologies offer the promise for more accurate genotyping, especially of complex genes, eventually leading to better gene function predictions. The field of mAb development, on the other hand, could benefit from technologies that can accelerate the mAb development process and are able to obtain antibody sequences from single B cells in a high-throughput manner. So, with this dissertation, some important steps have been taken on the road towards a more individualized and targeted approach to patient healthcare, helping to improve tomorrow's healthcare.