Functional study of the human genome

Li, Y.

Chapter 4

Summary

Summary

Understanding the meaning of the human genome codes is one of the keys to unlock the secrets of life. Despite having a fairly good grasp of the sequences of the human genome, we are still far from understanding the functions of most parts of the genome and their involvement in diseases. The application of the CRISPR-Cas9 genome-editing systems revolutionized the way to study the function of the genome, not only the coding genes but also the non-coding genome. In this thesis, multiple CRISPR screening systems were designed and used to study the transport of chemotherapeutic drugs and the functions of non-coding regulatory elements in distinct biological pathways.

In **Chapter 2**, in order to focus on drug transport and at the same time reduce the background interference from a genome-wide screening, custom CRISPR-knockout and CRISPR-activation libraries were designed and assembled to target all potential membrane-associated transporters. Taking advantage of the autofluorescence of doxorubicin, FACS sorting was used to enrich CRISPR-edited cells with either high or low drug accumulation during the screenings, in which drug uptake phenotype was directly measured. Such a strategy should avoid confounding factors or screening hits from other screening studies with similar aims that are based on cell survival as an indirect readout for drug transport. We showed that such screenings with focused libraries are very robust, as the same top hits were usually present in distinct populations indicating a similar potential drug transport function within the same screening, which in a way re-confirming their potential roles. Using this method, we identified previously known drug exporters such as ABCB1 and ABCG2. In addition, we identified a new doxorubicin importer gene SLC2A3 (GLUT3). We also realized that the CRISPR-activation screening complements the CRISPR-knockout system, which most of the research uses. For future screening efforts, it is better to combine the two screening systems to get a comprehensive identification of the potential factors involved in defined biological readouts.

In **Chapter 3**, we used an innovative dual-CRISPR system to study the functions of non-coding regulatory elements (NCREs) in their endogenous genomic environment. Using this method, we were able to study 4,047 UCEs in the human genome from UCNEbase, 1,527 in vivo-validated conserved enhancers from VISTA Enhancer Browser, and 13,539 potential K562-cell enhancers predicted by the ENCODE project, by deleting these NCREs one by one. We were able to identify many NCREs many affect cell fitness and drug responses. In addition, many of the NCREs actually have silencer (transcriptionally repressive) activity, in contrast to most commonly known enhancer elements. This is interesting and important because the roles of silencers in gene regulation in the whole genome were not well studied before. Our data also support the idea that silencers may exist broadly in the human genome and play important roles, similar to other well-studied NCREs such as enhancers and insulators. We were able to show that many of these NCREs regulate nearby genes, therefore affecting different biological pathways. One of the ultraconserved NCRE PAX6_Tarzan would affect the cardiomyocyte differentiation in an hESC model, suggesting many ultraconserved NCRE may exert their functions in a tissue-dependent context. We provide the broad research community with a new tool to study the functions of NCREs in different biological aspects, and it is expected many more important NCREs will be discovered in the future, which would not only be relevant to fundamental biology but also relevant to human diseases.