Minipig Genomics at Roche: **State of the Art**

Prof. Dr Ulrich Certa,[1] Dr Martin Ebeling,[2]
Dr Thomas Weiser[1] and Dr Thomas Singer[1]

Today, minipigs are becoming increasingly appreciated as animal models in drug research and predictive drug safety and toxicology. Their physiology and metabolism as well as responses to toxicants and medicines resemble the human situation in many aspects. By contrast with more established animal models like the mouse, rat or the cynomolgus monkey, only limited sequence information is available for many gene families of interest, like the cytochromes p450, drug transporters or cytokines. In addition, the choice of minipig-specific diagnostic assays and kits for clinical chemistry or histopathology is limited. As a result of the Macaca fascicularis genome project, we realised that knowledge of all mRNA and protein sequences is highly beneficial for the prediction of cross-reactivity of tests and reagents originally designed for use in humans. In addition, the deciphering of the cynomolgus genome enabled us to establish a variety of chip-based applications for genotyping, mRNA profiling, and copy-number variation analysis.

To decipher the genome of the Göttingen minipig, we based our work on the efforts of a large pig genome consortium consisting of 38 institutions which deciphered the farm-pig genome to advance research for animal production and breeding (http://piggenome.org/).

At the time of writing, an estimated 95–97% of the Göttingen minipig genome is known and assembled at Roche with about 18-fold sequence coverage. This genome version is already sufficient to deduce the minipig orthologues of virtually all human proteins, provided they are present in the minipig genome. When comparing conservation among protein-coding genes between human beings and the cynomolgus monkey, the overall sequence identity is about 93%. Based on available data, we estimate an average sequence similarity in the range of 85% for human and minipigs. This highlights the need for careful validation of available human assays and development of minipig-specific protein detection kits and reagents. The availability of the genome is an important step towards this goal because the cross-reactivity of human assays becomes predictable. TNF-a is considered a biomarker for systemic inflammation. If we wish to use a human ELISA kit to measure inflammatory responses in minipigs, for example, we can now align the TNF-a sequences of both species. In case the sequence identity is above 95%, the likelihood of cross-reactivity is high whilst identities below 85% are critical, especially when monoclonal antibodies are used for detection. Similarly, it is possible to predict whether a given therapeutic antibody cross-reacts with the minipig based on target identity.

Pharmacological activity as well as adverse events depend not only on sequence similarity but also on gene expression in the target tissue. For this purpose we applied next generation RNA sequencing to profile mRNA expression in heart, liver, spleen, blood, kidney and lung. A highly sensitive analytical method of this nature is required, especially for blood profiling. For example, genes that are only expressed by a small subpopulation of white blood cells have low-expression mRNA levels which are likely to escape detection by microarrays.

For more standard mRNA profiling applications like toxicogenomics, we are currently designing microarrays for mRNA profiling in collaboration with Roche-NimbleGen (Madison, USA). This technology is particularly suited for the design of prototype arrays since errors are easily corrected. Based on control hybridisations, we will assess performance and eliminate bad performing probes if necessary. An initially validated version of this array should be available in the autumn of this year, with probes for about 18,000 transcripts. Although the sensitivity of deep sequencing is considerably higher, microarrays are still the most robust tools for routine tissue expression analysis, backed up by well-established bioinformatics tools for data processing and analysis. Since pharmaceutical research mainly focuses on protein targets, microarrays will remain valuable tools for studying gene expression at different stages of preclinical drug development. It is likely that the pig genome community will also appreciate the availability of these arrays. Good compatibility for both organisms is more than likely based on the genome project and the fact that pig and minipig belong to the same species.

> *"With the availability of the genomes for mouse, rat, dog, minipig, and cynomolgus we can perform highly sophisticated tox-species selection which will not only reduce the number of animals required but also improve the prediction of drug safety in humans."*

In summary, minipig genomics support the decision-making drug-safety process on rational grounds. Sequence and gene expression information will largely inform the decision of whether the minipig is a responder species for the assessment of novel drugs. Consequently, "trial-and-error" experiments are avoided which will ultimately lead to a significant reduction in animal usage with a concomitant improvement of drug safety assessment. Genomics in general represents a very useful tool for refining and reducing animal experiments concordant with the 3R principles. With the availability of the genomes for mouse, rat, dog, minipig, and cynomolgus we can perform highly sophisticated tox-species selection which will not only reduce the number of animals required but also improve the prediction of drug safety in humans. Finally, novel minipig-specific tests will allow more reliable and specific measurement of safety parameters.

[1]Non-clinical Safety and [2]Translation Research Sciences, pRED
F. Hoffmann-La Roche AG, CH-4070 Basel, Switzerland