Blood and Biomarkers in Huntington's Disease

Mastrokolias, A.

Summary

Even though Huntington's disease (HD) and its most prominent symptoms were described in the middle of the 19th century, it was not until the last decades of the 20th century that a systematic analysis of the biology and genetics of the disease started. Great efforts by the scientific and patient community have raised disease awareness, reduced the disease stigma, allowed for a better and earlier recognition of potential symptoms and offered perspective for therapy development. The disease is known to exhibit geographic prevalence differences. The most recent data suggest that disease prevalence in Australia, North America and Western Europe has increased by approximately 15-20 % per decade for the second half of the 20th century. This increase accentuates the need to care for these patients and to identify symptomatic treatments and therapeutic approaches for the disease.

The work described in this thesis presents part of a framework that can be used to extract detailed disease biological information from peripheral tissue. This framework is based on the central dogma of biology "DNA to RNA to protein" and on a systems biology approach that aims to produce synergetic data whose disease pathological, prognostic and predictive value is greater than the sum of the individual experiment results. HD patients are often characterized by a multifaceted clinical profile, consisting of several symptoms and variable disease progression rates. Therefore, a systems approach such as the one described above is expected to be the most effective in identifying potential treatments and predictive biomarkers that will be most informative for the different patient subpopulations.

In Chapter 1, a detailed report of the disease background and symptoms is provided as well as current potential therapeutic approaches. In the last part of the chapter a thorough description is presented of the ideal properties of a disease biomarker. These properties are discussed in the context of the idiomorphic character of HD symptoms and, most importantly, the development of a blood biomarker reflecting brain pathology.

In Chapter 2 we investigate the depletion of the overabundant hemoglobin transcripts from next generation sequencing (NGS) gene expression peripheral blood samples. Even though technological breakthroughs such as NGS offer increased sensitivity and dynamic range the disproportionate number of hemoglobins in whole blood samples reduces the sequencing depth potential and the identification of the less abundant blood transcripts. We showed that by removing the hemoglobins, other types of transcripts such as lincRNAs, misc-RNAs and processed transcripts are increasingly and more reliably detected. To our knowledge this study was the first to describe the effects of such a depletion process on RNA sequencing data and especially in human blood samples. The various low abundance transcripts often play potentially important roles in disease in general which is also the case for Huntington's disease [1].

Chapter 3 describes the results of our whole genome gene expression profiling study employing NGS in peripheral whole blood samples. To date, our study constitutes the largest peripheral whole blood gene expression study; including such a substantial number of pre-symptomatic patients. Additionally, this study was the first HD study to have used next generation sequencing-based gene expression profiling, in contrast to hybridization microarrays which were previously the most commonly used platform. Our results confirmed for the first time blood gene expression changes found in previous studies, thus highlighting the potential of NGS to obtain disease relevant biomarkers. We showed that by employing appropriate statistical approaches, genes other than the most statistically significant ones can be used to establish disease progression biomarker panels. Such biomarker panels can consist of a combination of genes which, in addition to mere statistical significance, may provide complementary disease-specific information. These panels can potentially be refined to best reflect the pathological state in the different severity groups.

In Chapter 4 we performed an advanced integration of high throughput, large cohort data from blood and from more than one biological data classes. We integrated our targeted metabolomics results with our previously described NGS transcriptomics data from the same patients and controls. Metabolomics can reflect changes occurring downstream of the gene expression level and its deregulation that takes place in HD. Further-

more, targeted metabolomics has the advantage of enabling robust, reproducible and repeated measuring of the discovered changes in potential future longitudinal studies. By this multidisciplinary approach and employing community-driven bioinformatics research tools, we could link the changes in the HD patients' metabolite concentrations with specific subsets of the top deregulated genes of our gene expression dataset. Our integrative effort constitutes one of the first multi-level "omics" approaches, in the future to be enriched by additional biological data classes, thus providing a much-needed, more holistic description of HD disease pathology.

Finally, in Chapter 5 a critical evaluation was given of all the data produced and the conclusions drawn in the previous chapters. The knowledge gained by each of the above phases was assessed in the context of the novelty and applicability of the methodological approaches and platforms employed. In the penultimate part of this chapter the limitations of each of the above experiments were scrutinized with respect to the inherent technical limitations of the study designs and the analytic tools used. Finally, potential future directions and advancements associated with the study design, analytical components and bioinformatics tools used in each individual study were outlined.

References

1. Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J et al.: RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. PLoS One 2015, 10: e0143563.