Lars Juhl Jensen started his research career in Søren Brunak's group at the Technical University of Denmark (DTU), from where he in 2002 received the Ph.D. degree in bioinformatics for his work on non-homology based protein function prediction. During this time, he also developed methods for visualization of microbial genomes, pattern recognition in promoter regions, and microarray analysis. From 2003 to 2008, he was at the European Molecular Biology Laboratory (EMBL) where he worked on literature mining, integration of large-scale experimental datasets, and analysis of biological interaction networks. Since 2009, he has continued this line of research as a professor at the Novo Nordisk Foundation Center for Protein Research at the Panum Institute in Copenhagen and as a founder, owner and scientific advisor of Intomics A/S. He is a co-author of more than 170 scientific publications that have in total received more than 20,000 citations. He was awarded the Lundbeck Foundation Talent Prize in 2003, his work on cell-cycle research was named "Break-through of the Year" in 2006 by the magazine Ingeniøren, his work on text mining won the first prize in the "Elsevier Grand Challenge: Knowledge Enhancement in the Life Sciences" in 2009, and he was awarded the Lundbeck Foundation Prize for Young Scientists in 2010.

## Medical data and text mining: Linking diseases, drugs, and adverse reactions

Clinical data describing the phenotypes and treatment of patients is an underused data source that has much greater research potential than is currently realized. Mining of electronic health records (EHRs) has the potential for revealing unknown disease correlations and for improving post-approval monitoring of drugs for adverse drug reactions. In my presentation I will introduce the centralized Danish health registries and show how we use them for identification of temporal disease correlations and discovery of common diagnosis trajectories of patients. I will also describe how we perform text mining of the clinical narrative from electronic health records and use this for identification of new adverse reactions of drugs.