Genomics and bioinformatics for the study of infectious diseases

Björn Andersson, Department of Cell and Molecular Biology, Karolinska Institutet, Berzelius väg 35, 171 77 Stockholm. bjorn.andersson@ki.se, http://ki.se/ki/jsp/polopoly.jsp?d=11253&a=23134&cid=11255&l=en

My group was formed around a genome sequencing and analysis activity. The two main sequencing projects were initially the Trypanosoma cruzi genome project and a project for comparative human sequencing. I can offer project work involving both bioinformatics methods development and the application of these methods in the analysis of genomic data. There are two main areas:

1. The human virome

We have launched a large metagenomics effort to identify previously unknown human viruses. This is an urgent task, since the previous methods to identify viruses are very limited, and since many infectious diseases remain undiagnosed and many common diseases may in part be caused by viral infections. Our strategy has been shown to be effective and we have published two new viruses.

We use an enrichment protocol, followed by massive shotgun sequencing using the latest generation sequencers followed by bioinformatics analyses to identify viral sequences. A project worker will participate in improving the computational infrastructure for the project, the development of new methods for the identification of new distantly related viral sequences, and the actual analyses of new datasets from clinical samples.

2. Trypanosomatid genomics.

Trypanosoma cruzi is a protozoan parasite that is the causative agent of Chagas disease, which is a severe infectious disorder that affects millions of people in South and Central America. We were funded by the US National Institutes of Health to participate in the sequencing of the 100 Mb genome of this parasite. This made the project the largest genome sequencing project in the Nordic countries. The ultimate goals of the project are to learn more about the biology of the parasite and to find ways to treat parasitic disease. The complete genome sequence was published in a theme issue of Science in July 2005.

We have more recently initiated functional genomics studies in this parasite. We are focusing on the proteome of the parasite with the aim to characterize protein expression and localization in all stages of the lifecycle. We are also carrying out studies of specific proteins and protein complexes as well as comparative genomics of different parasite strains. The latter area will become more important during the coming year and it will yield important information regarding parasite pathogenicity. Project students will continue to develop methods for the analysis of difficult repeated regions of the genome and for comparative genomics, carry out such analyses and work with protein identification.

Selected publications

Wong G. K., Liu B., Wang J., Zhang Y., Yang X., Zhang Z., Meng Q., Zhou J., Li D., Zhang J., Ni P., Li S., Ran L., Li H., Zhang J., Li R., Li S., Zheng H., Lin W., Li G., Wang X., Zhao W., Li J., Ye C., Dai M., Ruan J., Zhou Y., Li Y., He X., Zhang Y., Wang J., Huang X., Tong W., Chen J., Ye J., Chen C., Wei N., Li G., Dong L., Lan F., Sun Y., Zhang Z., Yang Z., Yu Y., Huang Y., He D., Xi Y., Wei D., Qi Q., Li W., Shi

J., Wang M., Xie F., Wang J., Zhang X., Wang P., Zhao Y., Li N., Yang N., Dong W., Hu S., Zeng C., Zheng W., Hao B., Hillier L. W., Yang S. P., Warren W. C., Wilson R. K., Brandstrom M., Ellegren H., Crooijmans R. P., van der Poel J. J., Bovenhuis H., Groenen M. A., Ovcharenko I., Gordon L., Stubbs L., Lucas S., Glavina T., Aerts A., Kaiser P., Rothwell L., Young J. R., Rogers S., Walker B. A., van Hateren A., Kaufman J., Bumstead N., Lamont S. J., Zhou H., Hocking P. M., Morrice D., de Koning D. J., Law A., Bartley N., Burt D. W., Hunt H., Cheng H. H., Gunnarsson U., Wahlberg P., Andersson L., Kindlund E., Tammi M. T., Andersson B., Webber C., Ponting C. P., Overton I. M., Boardman P. E., Tang H., Hubbard S. J., Wilson S. A., Yu J., Wang J., Yang H.:
A polymorphism map for chicken with 2.8 million SNPs.

Nature 432: 717-722, 2004.

International Chicken Polymorphism Map Consortium

El-Sayed N. M., Myler P. J., Blandin G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E. A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D. C., Haas, B. J., Tran, A-N., Wortman, J. R., Alsmark, C. M., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J. M., Cerqueira, G., Creasy, T., Delcher, A. L., Djikeng, A., Embley, T. M., Hauser, C., Ivens, A. C., Kummerfeld, S., Leal, J., Nilsson, D., Peterson, J., Salzberg, S. L., Shallom, J., Silva, J. C., Sundaram, J., Westenberger, S., White, O., Melville, S. E., Donelson, J. E., Andersson, B., Stuart, K. D. and Hall, N.
Comparative genomics of the trypanosomatids.

Science 309: 404-409, 2005.

El-Sayed, N. M. A., Myler, P. J., Bartholomeu, D. C., Nilsson, D., Aggarwal, G., Tran, A-N, Ghedin, E., Worthey, E. A., Delcher, A. L., Blandin, G., Westenberger, S. J., Caler, E., Cerqueira, G., Branche, C., Haas, B., Anapuma, A., Arner, E., Åslund, L., Attipoe, P., Bontempi, E.
J., Bringaud, F., Burton, P., Cadag, E., Campbell, D. A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J. F., de Jong, P., Edwards, K., Englund, P. T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.
C. C., Gull, K., Horn, D., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M. J., Lorenzi, H., Louie, T., Machado, C. R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J. C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Martin Pentony, M., Pettersson, U., Pop, M., Ramirez, J. L., Rinta, J., Robertson, L., Salzberg, S. L., Sanchez, D. O., Seyler, A., Sharma, R., Simpson, A. J., Sisk, E., Tammi, M. T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P., Wickstead, B., Wortman, J., White, O., Fraser, C. M., Stuart, K. D. and Andersson, B.
The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas' disease.

Science 309: 409-415, 2005.

Allander, T., Tammi, M. T., Eriksson, M., Bjerkner, A., Tiveljung Lindell, A. and Andersson, B.

Cloning of a human parvovirus by molecular screening of respiratory tract samples.

Proc. Natl. Acad. Sci. USA 102: 12891-12896, 2005.

Arner, E., Tammi, M. T., Andersson, B.
DNP Trapper: A graphical workbench for shotgun sequence assembly and analysis of repeats.

BMC Bioinformatics, 7: 155, 2006.

--
Bjorn Andersson Ph.D.

Professor of Genome Analysis
Department of Cell- and Molecular Biology
Programme for Genomics and Bioinformatics
Karolinska Institutet
Berzelius vag 35
S-171 77 Stockholm
Sweden

phone +46-8-5248 3987
fax +46-8-323950
mobile: +46-70-3553987
http://cruzi.cgb.ki.se/