Current Status and Call for Sequencing Requests of *Idd* regions in the NOD mouse

The National Institute of Allergy and Infectious Diseases (NIAID), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), and the Juvenile Diabetes Research Foundation International (JDRF) are pleased to announce the availability of resources for sequencing *Idd* regions in the NOD mouse:

The Wellcome Trust Sanger Institute, Cambridge, UK, has generated the sequences from the ends of BACs from two independent NOD mouse libraries: a library from the Diabetes and Inflammation Laboratory at the University of Cambridge, which was developed from the NOD/MrkTac mouse strain (NOD DIL); and a NOD BAC library (CHORI-29; Pieter de Jong, Research Genetics), which was developed from the NOD/LtJ mouse strain.

The NOD mouse clone mapping to the mouse genome sequence is available through DAS tracks in Ensembl Mouse Contig View and Cytoview. The current Ensembl release shows the clones mapped to NCBI M33; the clones will be re-mapped as subsequent builds of the mouse genome sequence are released. The data can be accessed from the entry point: (http://www.ensembl.org/Mus musculus/).

A summary of the sequencing and mapping of both libraries to NCBI M33 (current as of 11/1/04) is shown:

Library	DIL NOD	CHORI-29
Clones end sequenced	180217	93023
Successful reads	332535 (92.2%)	170159 (91.1%)
No. of reads with alignment to C57B6/J sequence	279179 (84.0%)	158973 (93.4%)
No. aligned reads with unique match	225410 (80.7%)	109307 (68.7%)
No. aligned reads with probable best match	28120 (10.0%)	22374 (14.1%)
No. reads with multiple matches	25649 (9.2%)	27292 (17.2%)
Clones with paired end sequences	150902 (83.7%)	75046 (80.7%)
Clones positioned uniquely using read pairs	103711 (68.7%)	60217 (80.2%)
Clones with multiple positions	1533 (1%)	1058 (1.4%)
Average depth of genome coverage	6-fold	5-fold

The distribution of mapped sequence reads is now relatively even across the genome, with the exception of the Y chromosome, which is not represented in the DIL NOD mouse library. As the reference mouse genome sequence continues to be finished, we anticipate that more DIL and CHORI-29 clones will be placed on the genome.

All interested investigators are invited to propose <3 Mbp *Idd* intervals for full, finished sequencing or candidate genes for gene-based genomic sequencing.

Written proposals will be considered on a case-by-case basis by the External Advisory Committee for this project.

Proposals with the following will be considered:

1) Intervals less than 3 Mbp defined by NOD congenic strains, with evidence showing that the interval alters the frequency of diabetes. Information on independent or complementary congenic strains that support the definition of the *Idd* interval should be included, if available. If approved, the entire *Idd* region will be sequenced.

2) Candidate genes mapping within a genetic interval defined by a genome scan for diabetes (or other fundamentally-related phenotypes such as peripheral or central tolerance). The genome scan must involve the NOD background and evidence of expression differences should be included. In this case, one or two BAC clones encompassing the candidate gene will be sequenced rather than the entire genetic interval under the linkage curve. More than one candidate gene in a linked region can be requested for consideration.

3) Candidate genes identified by expression differences in intervals less than 10 Mbp defined by NOD congenic strains, with evidence showing that the interval alters the frequency of diabetes. Information on independent or complementary congenic strains that support the definition of the *Idd* interval should be included, if available. In this case, one or two BAC clones encompassing the proposed candidate gene will be sequenced rather than the entire genetic interval defined by the congenic strain. More than one candidate gene in a region defined by such a congenic strain can be requested for consideration.

4) Candidate genes identified from knock-out experiments on a NOD background. The knock-out must have an effect on the frequency of diabetes. If the gene was disrupted on a non-NOD genome such as 129 or B6, the extent of the linked region that exists in the KO mice must be documented. Evidence of natural allelic variation in gene expression between the NOD strain and at least one other mouse strain (preferably B6, since this is the reference strain in Mouse Ensembl) should be included in the proposal to be considered.

Proposals may be submitted at any time, via email to: John Ridge (jridge@niaid.nih.gov) or Beena Akolkar (akolkarb@extra.niddk.nih.gov).

Proposals will be considered by the External Advisory Committee for this project. Investigators will be notified by email, and will receive instructions on how to proceed. ***All sequence data will be released according to Bermuda rules: there will be no provision for any exclusive access to the sequence data.***

Two regions are presently in the sequencing pipeline: *Idd 4.1* (chr 11), and *Idd 18* (chr 3).

The Sanger Institute is constructing a Web site that will provide users with up to date information about the status of the sequencing projects and the progress of individual clones in the sequencing pipeline. This site will be active December 2004.

For more information, or in case of questions, please contact John Ridge (jridge@niaid.nih.gov), Beena Akolkar (akolkarb@extra.niddk.nih.gov), or Marie Nierras (cnierras@jdrf.org).