User Guide for DNAFORM Clone Search Engine

Document Version: 3.0

Dated from: 1 October 2010

DNAFORM *Precision Gene Technologies*

List of Content:

DNAFORM *Precision Gene Technologies*

1. Introduction

Welcome to the DNAFORM Clone Search Engine! The DNAFORM Clone Search Engine was developed to provide our customers with better tools to search our genomic resources for the clones of interest. In addition, reference information to resources purchased from us can be found to help you getting the best out your genomic resources.

We made best efforts to use most up-to-date reference information on each of our clones by linking them to the gene annotations at NCBI. For easy use, entry pages are furnished with short descriptions on how to enter your information. All information is provided in English and Japanese. This User Guide provides additional information on how to search the database, the retrievable clone information, and how to place orders on our online shop.

2. Database Content

By clicking on the link for "cDNA Clones Content" and "shRNA Clones Contents" on the opening page an information page will be opened that informs you about the content of the database and the clones that can be searched. Presently the DNAFORM Clone Search Engine supports all full-length sequenced RIKEN FANTOM Clones of the FANTOM 3 Set, RIKEN human cDNA and human Gateway entry ORF clones, the RIKEN honeybee cDNA collection, the human and mouse Mammalian Gene Collection (MGC) clone sets, the human and mouse ORF clones from the ORFeome Collaboration Project, and human shRNA expression constructs from The Netherlands Cancer Institute (NKI) in the Netherlands. DNAFORM will add new information to the database when ever new clones are released. Refer also to the public announcements on our homepage for new genomic resources offered by DNAFORM.
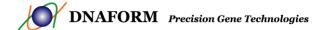
For more information on those clone sets having project databases refer to the Internet pages of the originators under the following links:

FANTOM Project: http://fantom3.gsc.riken.jp/

Mammalian Gene Collection (MGC): http://mgc.nci.nih.gov/

ORFeome Collaboration: http://www.orfeomecollaboration.org/

shRNA clones: http://www.nki.nl/Research/

NOTE, whereas the MGC, RIKEN FANTOM clone collections,RIKEN honeybee cDNA and the RIKEN human cDNA clone comprise original cDNAs including untranslated regions and non-coding RNAs (RIKEN FANTOM only), the ORF clones from the ORFeome Collaboration Project and the RIKEN human Gateway entry ORF clones contain only the protein coding regions (Open Reading Frame or ORF) flanked by recombination sites for the Invitrogen Gateway System. Consider ORF clones for direct expression of proteins of interest in any Gateway Expression Vector. Refer to the homepage from Invitrogen for more information on the Gateway System:

http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/Gateway-Cloning.html
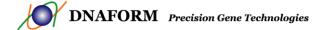
The shRNA resource has been designed for screening experiments. Commonly up to 3 different shRNAs are available for each target gene to assure that at least one of them is functional. It is recommended to purchase always all shRNA constructs available for your target gene. NOTE FURTHER, DNAFORM can also provide additional cDNA clones, including EST clones, and genomic clones that cannot yet be found in the DNAFORM Clone Search Engine. For more information on our other products and services, refer to our homepage under http://www.dnaform.jp or contact us directly under order@dnaform.jp.

3.   Navigation

All pages of the DNAFORM Clone Search Engine provide the following functions to guide you to other DNAFORM pages with additional information on our services and products:

- Press "Home" to return to the DNAFORM homepage.

- Press "cDNA Library" to get information on our full-length cDNA library services.

- Press "DeepCAGE Service" to get information on our DeepCAGE services for expression profiling and genome annotation.

- Press "Clone Distribution" to get information on our clone distribution services and resources.

You can click on the DNAFORM Logo in the upper left corner on each page within the DNAFORM Clone Search Engine to return to the opening page of the DNAFORM Clone Search Engine.

All text written in "blue" has a hyperlink that opens a new page with additional information on individual clones, vector maps, or entries in public databases.

On the opening page you will further find links to:

- About Genomic Resources: Provides information about clone distribution business.

- Disclaimer: On rights to use DNAFORM Clone Search Engine.

- Feedback: Contact address to provide feedback on DNAFORM Clone Search Engine.
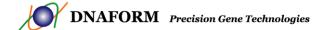
4.   Search Functions

The DNAFORM Clone Search Engine has independent search windows for "cDNA clones" and "shRNA clones". Clicking on the headers will list the data that can be search by the available tools. For searches for "cDNA clones" three different Search Functions can be used: "Search by clone or gene ID", "Search by plate address", and "Search by BLAST nucleotide sequence". For searches for "shRNA clones" two different Search Functions can be used: "Search by clone or gene ID", and "Search by nucleotide sequence". Each of these functions offers specific solutions to meet with your requirements.

5.   cDNA clone search by clone or gene ID

While building the clone collections, each clone got one or more "clone ID" as unique identifier(s) to track the clones later in public databases or publications by clone users. The "Search by clone or gene ID" function allows you to use those clone IDs to look for specific clones.

In addition, a unique accession number was assigned to every clone related sequence submitted to GenBank or other public databases. The DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI are members of the International Nucleotide Sequence Database Collaboration using the same accession numbers. The "Search by clone or gene ID" function allows you to find directly the matching clone for an accession number. NOTE, that only accession numbers for clone derived sequences were incorporated in the database. Therefore you cannot use accession numbers for computational generated reference sequences as for instance found in NCBI.

For gene related information, use the official "gene symbol" or "gene ID" to search the DNAFORM Clone Search Engine. For more information on "gene symbols" and "gene IDs" refer to Entrez Gene at NCBI under: http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene.

Refer to the "Search by BLAST nucleotide sequence" if you cannot find your clone of interest by using a clone ID, accession number, gene symbol, or gene ID.
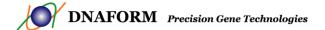
6.   cDNA clone search by plate address

The "Search by plate address" function provides an information tool to our clone customers, who have purchased clone plates from DNAFORM. Use the dropdown menus to enter coordinates for the plates or individual position on a plate to retrieve clone related annotation information. Please indicate under "Collection" in which clone set (FANTOM, human cDNA, human ORF entry, honeybee cDNA, MGC, or ORF) and "Collection" (sets of remarried plates within "Libraries") you want to search for a certain plate. For the MGC and ORF clone sets you must further select the "Species" to distinguish between "human" and "mouse" clones.

7.   cDNA clone search by BLAST nucleotide sequence

The "Search by BLAST nucleotide sequence" function offers you a convenient and more flexible way to look for clones if you do not have any clone ID, accession number, Gene Symbol, or Gene ID at hand, you want to search for related clones (e.g. splice variants), or you have your own sequence data and want to indentify related clones. Enter your nucleotide sequence into the search box and press "search" to run BLAST under default conditions. This will commonly provide you with reliable and high quality alignments and hits in the DNAFORM Clone Search Engine. If you wish to change the default settings to run BLAST, the "Search by BLAST nucleotide sequence" function allows you change parameters for the "minimum match length" and "minimum match quality" prior to running your search.

8.   shRNA clone search by clone and gene ID

To find suitable shRNA clones use the official "gene symbol" or "gene ID" to search the DNAFORM Clone Search Engine. For more information on "gene symbols" and "gene IDs" refer to Entrez Gene at NCBI under: http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene.

DNAFORM *Precision Gene Technologies*

Accession number searches are also possible to identify target gene for shRNAs.

9.   shRNA clone search by nucleotide sequence

The DNAFORM Clone Search Engine allows you to use sequences e.g. from cDNA clones or transcript information in the public databases to find relevant shRNA clones by a nucleotide sequence. Note, that shRNAs have only about 19 nucleotides. Therefore this search function will only show hits with perfect sequence matches.

10.   Search Results for cDNA Clones

The DNAFORM Clone Search Engine will provide you with comprehensive information on each clone found by your search criteria. Make use of this information to select the best clones for your order or to get more information on clones purchased from DNAFORM.
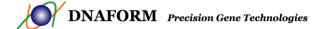
10.1   Select

"Select" allows you to mark the clones you want to purchase from DNAFORM. Continue with the "Preview order" function after you have selected the clones for your order.

10.2   ID

"ID" lists the unique clone ID assigned to each clone. Clicking on the clone ID (marked in blue) will open a new window with clone-specific information. On this window you can find information regarding:

-    Clone collection and plate position,

-    Vector information (click on vector map to access vector map),

-    Clone related information,

-    Gene related information/annotation information from NCBI,

-    Nucleotide sequence.

-    For human clones, this window has the option to "Find shRNA products for this Clone". Clicking on this option will run a search to find shRNA clones within our collection for the same gene.

10.3   Collection

"Collection" tells to which clone collection a clone belongs to. FANTOM, human cDNA, human ORF entry, honeybee cDNA, MGC, or ORF clone collection a clone belongs. NOTE the special features of ORF clones if your search indentified ORF clones.

10.4   Species

"Species" indicates from which species the clone had been obtained. Depending on your search criteria, you may find clones from different species.

10.5   Host Strain

"Host Strain" indicates the E. coli strain in which a clone is maintained.

10.6   Vector Map

"Vector Map" name the vector used for a given clone. Click onto the vector name to access the vector map.
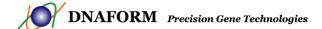
10.7   Accession

"Accession" lists the accession number under which clone derived sequences have been deposited at GenBank. Clicking on the accession number (marked in blue) will open a new window with the entry in NCBI under this accession number.

NOTE there may be multiple accession numbers for some cDNA clones related to partial (EST) sequences and full-length cDNA sequences.

10.8   Gene Accession

"Gene Accession" lists Gene Accession numbers where such an identifier had been originally assigned to a cDNA clone. Clicking on the Gene Accession number (marked in blue) will open a new window with the entry in NCBI under this Gene Accession number. NOTE for most clones no Gene Accession numbers are available, and you may wish to use "Gene Entrez" or "Gene Symbol" to get more information on the related gene.

10.9   Gene Entrez

"Gene Entrez" lists the Gene ID assigned by NCBI within Gene Entrez. Clicking on the Gene ID (marked in blue) will open a new window with the entry in NCBI under this Gene ID number. NCBI provides comprehensive information on each gene including cross references to other information in their databases.

10. 10   Gene Symbol

"Gene Symbol" lists Gene Symbols within NCBI. Gene Symbols are unique abbreviations of gene names assigned by the HUGO Gene Nomenclature Committee after a gene has been identified. Clicking on the Gene Symbol (marked in blue) will open a new window with a publication record on this particular gene.

10.11   Plate Set

"Plate Set" tells to which plate set within the FANTOM, human cDNA, human ORF entry, honeybee cDNA, MGC, or ORF clone collections a clone belongs. This information is needed to find a clone on a specific plate.

10.12   Plate

"Plate" tells on which plate within a "Collection" a clone is located. This information is needed to find a clone on a specific plate.

10.13   Row and Column

"Row" and "Column" tell the position of a clone on a specific "Plate". This information is needed to find a clone on a plate.

10.14   Query Length (BLAST Search only)

"Query Length" indicates the number of nucleotides entered for the BLAST search.

10.15   Hit Length (BLAST Search only)

"Hit Length" indicates the number of nucleotides matching with the aligned sequence.

10.16   Hit Quality (BLAST Search only)

"Hit Quality" gives the percentage of matching nucleotides within the aligned sequence.

DNAFORM *Precision Gene Technologies*

11. Search Results for shRNA Clones

The DNAFORM Clone Search Engine will provide you with comprehensive information on each clone found by your search criteria. Make use of this information to select the best clones for your order or to get more information on clones purchased from DNAFORM.

11.1 Select

"Select" allows you to mark the clones you want to purchase from DNAFORM. Continue with the "Preview order" function after you have selected the clones for your order.
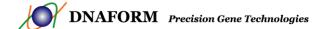
11.2 ID

"ID" lists the unique clone ID assigned to each clone. Clicking on the clone ID (marked in blue) will open a new window with clone-specific information. On this window you can find information regarding:

- Clone product,

- shRNA sequence,

- Species,

- Accession (link into Genbank),

- Symbol (link to publications on PubMed),

- Entrez,

- HGNC,

- Description.

- This window further has the option to "Find cDNA clones for this Product". Clicking on this option will run a search to find cDNA clones within our collection for the same gene.

11.3 Sequence

"Sequence" lists the sequence of the shRNA clone.

11.4 Symbol

"Gene Symbol" lists Gene Symbols within NCBI. Gene Symbols are unique abbreviations of gene names assigned by the HUGO Gene Nomenclature Committee after a gene has been identified. Clicking on the Gene Symbol (marked in blue) will open a new window with a publication record on this particular gene.

11.5   Accession

"Accession" lists the accession number for the reference sequence deposited at GenBank. Clicking on the accession number (marked in blue) will open a new window with the entry in NCBI under this accession number.

11.6   Entrez

"Entrez" lists the Gene ID assigned by NCBI within Gene Entrez. Clicking on the Gene ID (marked in blue) will open a new window with the entry in NCBI under this Gene ID number. NCBI provides comprehensive information on each gene including cross references to other information in their databases.

11.7   HGNC

"HGNC" provides a gene-specific link to the HGNC Database and resources for the human genome.
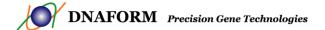
11.8   Species

"Species" indicates from which species the shRNA clone had been designed (presently only human).

12.   Preview Order

The "Preview order" function allows you to review the clones you have selected for purchase. Besides just listing the clones, it allows you again to retrieve clone related information. Check the boxes under "Select" to choose the clones of your interest. Press "Preview order" to precede your order, press "Search again" to start a new clone search.

13.   Order

The Order page lists your selected clones. Confirm that the box under "Select" is checked. Press "Cancel" to return to the Preview page, press "Order" to continue the ordering process. Confirm your order one more time and you will be forwarded to our online shop and you are asked to

provide the necessary information to process your order. NOTE, all clones are offered under the terms of a Good Faith or Sublicensing Agreement. DNAFORM is obliged to execute those agreements on behalf of the originators of the clone collections.

14. Disclaimer

The information in our database has been taken from the public data resources of the RIKEN FANTOM, MGC, ORFeome Collaboration, and NKI projects. Although the information has been reviewed by us while building our database, K.K. DNAFORM cannot exclude that some of the information may contain mistakes. Therefore all the information in our database is considered to be of experimental nature, and K.K. DNAFORM does not take any responsibility for the correctness of the information.

The DNAFORM Clone Search Engine is maintained by K.K. DNAFORM to provide information to our customers and Internet users. The provided information is "as is", and K.K.DNAFORM gives no warranties on the accuracy or completeness of the information in our databases. K.K.DNAFORM shall not be liable for damages of any kind whether direct, consequential or punitive arising from or related to information provided by the DNAFORM Clone Search Engine.

K.K. DNAFORM takes no responsibility for any of the information and content offered on pages maintained by other parties and linked to the DNAFORM Clone Search Engine. None of those homepages is owned or otherwise related to K.K. DNAFORM.

K.K. DNAFORM keeps the right to revise, update or change the content of the DNAFORM Clone Search Engine at any time. K.K. DNAFORM will make best efforts to maintain and update the information in the DNAFORM Clone Search Engine.

K.K. DNAFORM will enforce its intellectual property rights on product names, trademarks, trade secrets and copyrights as protected by law. The entire content of our web pages including any text, graph or photo is the property of K.K. DNAFORM and shall not be modified, reposted, downloaded or distributed without the expressed written permission of K.K. DNAFORM.

15. Your Feedback

The DNAFORM Clone Search Engine was developed to serve our customer. It is your feedback that will help us to improve the system, identify mistakes, and to provide you with the tools and

information you need. Please provide us with your feedback, comments, and suggestion under techinfo@dnaform.jp.