**Nematode.net: A Community Resource for Nematologists**

This document is a supplement to the presentation “**Nematode.net: A Community Resource for Nematologists**”. It’s indexed by slide number and provides URLs & instructions to help you try out the web tools & views being presented in the talk.

To get the most out of this talk its best to follow along in the live website and try things out for yourself.

Due to the length of some of the URLs, I suggest having this document open electronically so you can copy-paste directly into your browser if needed.

Presenter: John Martin jmartin@genome.wustl.edu

Website: Nematode.net [http://www.nematode.net](http://www.nematode.net/NN3_frontpage.cgi)

Citation:

[Martin J., Abubucker S., Heizer E., Taylor C.M. and Mitreva M. (2011) **Nematode.net update 2011: addition of data sets and tools featuring next-generation sequencing data** *Nucleic Acids Research* first published online December 1, 2011 doi:10.1093/nar/gkr1194](http://nar.oxfordjournals.org/content/early/2011/12/01/nar.gkr1194.abstract)

**Slides 1-3:**

Introduction, Outline of talk & an Introduction to Data Navigation in Nematode.net.

URL:

nematode.net

**Slides 4-5:**

These slides are talking about the Species Hubs, which are organism-based portals into Nematode.net resources.

Notice the expandable menu of species that takes up the left-hand side of the page. Open the “Human Parasitic Nematodes” category, and from the list select Brugia malayi

URL:

nematode.net (start here)

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=speciestable&subnav_selection=Brugia_malayi> (species hub for B.malayi)

**Slides 6-7:**

Slides introducing the Data-mining and Comparative Genomics entry portals.

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nemapath> (NemaPath)

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=search_splash_page> (HelmCoP)

**Slide 8:**

This slide describes Nematode.net’s home page.

URL:

nematode.net

**Slide 9:**

This slide describes the contents of the ‘Home’ menu along the top navigation bar.

URL:

nematode.net

**Slide 10:**

This slide describes the content of the species hub pages.

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=speciestable&subnav_selection=Brugia_malayi> (species hub for B.malayi)

**Slide 11-13:**

These slides describe the **NemaGene Cluster Search** tool. *Slide 12* shows the primary interface, which is located at

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemagene_cluster_search>

Using this interface requires:

1. selecting a ‘Search Type’ (\*do this first, since the page will reload on selection)
2. selecting a ‘Species Database’
3. entering a search term

The page suggests some examples of each search type. To duplicate the results shown in the slides, select search type “Stage”, species database “Brugia malayi”, and from the stage dropbox select “L3”.

This search takes you to an intermediate page, shown on *slide 13*, listing all the Brugia malayi contigs that match your search (in this case “L3”) organized by cluster.

Scroll down the list and select the contig “BM00279” to reproduce the information display shown on *slide 13*. It may be easiest to use your browser’s text search to find BM00279, which is a member of the cluster BM00070.cl. You can also just pick any contig or cluster if you just want to see the information view page. The BM00279 information view can be seen at

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=detail_report&species_for_detail_report=Brugia%20malayi&type_for_detail_report=contig&name_for_detail_report=BM00279&stage_for_detail_report=L3> (direct URL to BM00279 information page)

Notice that this is just a ‘post’, so it’s easily possible to bypass the search interface if you know the name of the contig or cluster you want to view.

**Slide 14-17:**

These slides describe the **NemaGene Gene Table**. This resource is accessed by selecting a species of interest from the table at:

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemagene_gene_tables>

(only B.malayi is currently available). After clicking that link you’ll be sitting in the gene table for B.malayi being shown in *slide 15*. The direct URL to that location is

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=bmalayi_gene_table> (B.malayi gene table)

The table legend (being shown in *slide 16*) is available from the gene table by clicking the words “(Table Legend)” that appears to the right of the table title. Keep in mind that the annotations described in the legend will change for different projects. To directly access the table legend for B.malayi go to

URL:

<http://www.nematode.net/lib/nemagene_gene_tables/bmalayi_table_legend.html> (B.malayi gene table legend)

**Slides 18-27:**

This section of the talk describes **NemaPath**

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nemapath> (NemaPath start page)

To experiment with NemaPath, the first thing to do is to select which view you want to use by clicking one of the links near the top of the page:

**Species-specific NemaPath comparisons** - for exploring a single species, or building comparative views between 2 species

**Clade-specific NemaPath comparisons** – this view divides all the data in NemaPath by clade, and builds comparative pathway maps showing the presence of absence of clades per node (enzyme) in the selected pathway

**Host-specific NemaPath comparisons** – this view divides all the data into 3 broad classes of host organism: i) animal, ii) plant & iii) free-living. Comparative maps are drawn along those classes

**Stage-specific NemaPath viewer** – This view is currently in its pilot stage, with only a single pair of stages for a single organism having been loaded. Click this link to view comparative maps between the L3 and Adult stage of Necator americanus.

*Slide 20* shows the species selection menu arrived at by choosing Species-specific NemaPath comparisons. The direct URL is

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=species_nemapath> (Species-specific NemaPath comparison baseline species selection menu)

*Slide 21* shows the mapping cutoff selection page, used to filter the set of alignments that will be used to populate the pathway maps you view. Setting small (stringent) evalues (such as 1e-20) limits false positives at the expense of sensitivity. Weak evalues (such as 1e-03) increase sensitivity, but can introduce questionable mappings. In the case shown, B.malayi has a relatively large amount of data mapped, so it’s safe to select a fairly stringent evalue cutoff.

The cutoff selection page displayed on *slide 21* is normally accessed simply by clicking on a species link (as shown in *slide 20*). But a direct link to this page would be

URL:

<http://www.nematode.net/lib/nemapath/cgi-bin/KEGGscan_hit_distribution.cgi?species_selection=Brugia%20malayi%20genes> (cutoff selection page for B.malayi in NemaPath)

Note that after choosing and submitting a cutoff on this page, you will be presented with a confirmation page. Simply click submit again to continue on to the pathway selection menu.

*Slide 22* displays the pathway selection menu. From here choose the pathway you want to view from the dropdown menus. Note that this page cannot be accessed by direct URL. Simply open one of the menus and make your selection, the selected pathway view will load on selection.

*Slide 23* shows the baseline view for Brugia malayi in the ‘Alanine, aspartate and glutamate metabolism’ pathway. Again this page is dynamically generated and cannot be reached by direct URL.

\***NOTE**: script based re-direction into specific NemaPath views is possible based on KO mappings. If you are interested in integrating your web-based, KO annotated data with NemaPath drop me an email (John Martin – jmartin@genome.wustl.edu).

You can interact with this page in a number of ways:

1. clicking on any unpopulated (gray) enzyme node will open the KEGG orthology entry for that EC number.
2. Mousing over any populated node (green or other colors depending on whether you’ve loaded a comparative species or not) summons a popup. Links in that popup lead to appropriate resources . Note that the popups only display up to the top 10 hits, but you can access a complete list of hits using the link near the top of the popup box.
3. The 2 drop menus in the upper right corner of the NemaPath view provide you the ability to switch between pathways & to select a 2nd species for comparison.

*Slide 24* shows a comparative view between B.malayi & C.elegans. This page is navigated in the same fashion as before, the only difference being that now the popup information boxes show hits from both species in their respective colors (yellow for B.malayi & blue for C.elegans in the case on *slide 24*). You can freely choose the comparative species, but to select a new baseline species (displayed as the ‘Current Species’ in the upper right hand corner display), you need to go back to the species index and start again. There is a ‘species index’ link available in this view in the upper left hand section of the page, under the Nematode.net logo.

*Slide 25* shows the clade-based NemaPath comparisons. This page works in much the same fashion as the species-specific views, with the notable exception that in this view all hosted organisms in NemaPath are being used to populate the pathway map. Thus there is no menu for selecting species in this view.

*Slide 26* shows the host-specific NemaPath comparisons. Again this view is similar to the previous views except that all organisms, organized by host, are being displayed. Again no species selection is available in this view.

*Slide 27* shows the stage-specific NemaPath view. For this view, only the L3 & Adult stage data for Necator americanus has been loaded. The interactive portions of the page are the same as before. This will most likely not be the final form of this view, which we still consider to be in the beta stage.

**Slides 28-33:**

These slides are describing the **HelmCoP** analysis interface.

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=search_splash_page> (HelmCoP information page)

To use HelmCoP you first need to decide whether you want to query the database with the atomic unit being the gene or the ortholog. HelmCoP’s **Gene Search** interface is most useful if you either have a specific gene in mind, or want to mine information for all genes from one of the HelmCoP member organisms. The **Ortholog Search** is more useful if you want to setup a broad search, and want to be sure you catch all potential members defined by that search (since genes not fully annotated, but orthologous with something that does match your search will get pulled into the result).

**NOTE**: Some queries can be very slow to run. There is a query manager will try to catch queries that we anticipate will take too long (leading to a server timeout), but some queries that do work can sometimes take 3-4 minutes to complete.

*Slide 29* shows the species selection portion of the **HelmCoP Gene Search** interface

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=gene_search> (HelmCoP Gene Search interface)

Selecting multiple species here simply widens the initial pool of genes you are querying before applying filters.

*Slide 30* shows the species selection portion of the **HelmCoP Ortholog** Search interface

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=gene_search> (HelmCoP Ortholog Search interface)

Unlike with the Gene Search interface, including & excluding species affect your initial pool of pre-filter orthologous groups of genes by requiring each reported orthologous group to have at least 1 member from the included species. Then after the set of orthologs with all the included species are built, the exclude species are applied, and any ortholog in the list with at least 1 member from an exclude organism will be excluded. Exclusions take priority over inclusions.

*Slide 31* illustrates the filters you can set and the output information you can request. Both the Gene Search & Ortholog Search interface include this same bank of filters & output options.

With regards to the complexity of the query being made, and the time will take to run, applying filters generally *limits* the pool of genes or orthologs being reported on, and tend to make faster queries. While requesting more output options tends to make queries run slower as more tables are navigated to find information. The query manager does a fairly good job of catching extremely slow queries, but it can be helpful to think about this while choosing filters & output options.

*Slide 32* illustrates the HTML output table generated. The exact format varies depending on the outputs you request, but a table header row will always be provided. It’s very important to remember that HTML output is limited to 20,000 rows. Any additional rows are truncated from the HTML output. Full results for your query are best viewed by using the “**Full Results Download**” button. Note that clicking that button actually re-launches your query without the 20,000 row limit applied. Because its re-running your query, building the text output can be slow, but its typically much faster than the first run, since views have already been made based on your request. Once the query is finished you’re provided a link to the output. Right-click-save to download the data.

*Slide 33* shows the HelmCoP WU-BLAST v2.0 interface

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=helmcop_blast> (HelmCoP BLAST interface)

For additional information on HelmCoP, see the FAQ available at

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=helmcop_faq> (HelmCoP FAQ)

**Slides 34-37:**

These slides describe the **NemaGene BLAST service**

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemablast> (NemaBLAST front page)

To run WU-BLAST v2.0 against our NemaGene data, you first need to decide whether you want to align against Sanger **EST reads grouped by library**  or by **Transcript contigs, isotigs & genes**.

The **EST reads grouped by library** option limits the database to only Sanger EST reads that make up the EST contig portion of NemaGene. While this does not span the breadth of data in NemaGene, because you can search by library, this does offer some limited ability to select libraries for specific stages.

*Slide 35* shows the 2 forms you will move through when setting up your blast vs. EST reads. Notice on the 2nd form on the right side of the page, individual libraries can be selected or deselected (by default all libraries under your chosen organisms are selected). For additional details about libraries you can look them up in Nematode.net’s “Library Details” page

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=home&subnav_selection=library_details> (Library Details page)

*Slide 36* shows the form you arrive at if you select to map against **Transcript contigs, isotigs & genes**. This set represents a full view of all the data hosted in NemaGene, and allows you to choose the organisms against which you want to search.

**Slides 38-40:**

These slides describe the **NemaSNP** resource. Start by selecting a project

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=pnema_snpdb> (NemaSNP start page)

Selecting a project link takes you to a table of contents for that project (shown in *slide 39*). In the example case (which is the only case currently available on Nematode.net), the references are transcript assembly contigs, and the table provides information on how many field vs. inbred snp loci were found on each contig

URL:

<http://www.nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=pnema_snpdb_tcirc> (NemaSNP T.circ. project table of contents)

*Slide 40* shows a typical GBrowse view

URL:

<http://gmod.wustl.edu/cgi-bin/gb2/gbrowse/nemasnp/?name=contig02609> (NemaSNP GBrowse view for T.circ. transcript contig02609)

Note that in addition to using the NemaSNP table of contents, you can navigate using the GBrowse interface directly. Simply enter a contig name into the search window to jump directly to another transcript contig.

**Slides 41-42:**

These slides describe the **NemaBrowse** repository for nematode genomes

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemabrowser> (entry page for NemaBrowse)

Selecting a genome takes you to a table of contents listing each predicted gene. In *slide 42* the example is showing the T.spiralis annotations. The GBrowse view in this slide is looking at TSP\_00001 at

URL:

<http://gmod.wustl.edu/cgi-bin/gb2/gbrowse/t_spiralis_geneset/?name=SCTG0%3A2407..3194> (NemaBrowse GBrowse view for T.spiralis gene TSP\_00001)

Normal GBrowse navigation can be used, or you can drop back to the table of contents to move between annotations.

**Slides 43-45:**

These pages go over the **AmiGO viewer**, which hosts Nematode.net’s GO annotations for NemaGene data. The on-ramp into these annotations is at

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=go_associations> (GO Associations start page)

From there select an organism to explore and click the “GO” link. This will bring you to an expandable table listing the number of annotations for the current organism under each GO term (shown on *slide 44*). Notice the little pie chart image sitting next to each branch you have expanded. Clicking on that image will open a supplemental page with a breakdown of membership at that term’s level and beneath (all its child terms).

Clicking on the term itself (or anywhere on the line other than the expand/collapse (+/-) symbol) will open a detailed view of that term and all the annotations for the current species assigned to that term (shown on *slide 45*). Towards the bottom of that view is a listing of the transcripts assigned to that term. This view can also be accessed directly by entering a GO term into the “Search GO” box on the top AmiGO page. Terms must be entered in the form “GO:0016032”, and the search type must be set to “Terms”.

**Slides 46-48:**

This set of slides describes our **NemFam** interface

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nematode_protein_families> (NemFam front page)

The NemFams currently hosted and accessible via this interface are based on Sanger EST contigs. Additional NemFam collections have been made, but only the original 2009 dataset is explorable via the provided NemFam tool on Nematode.net.

You can jump into the data in several ways. If you know a NemFam id, you can simply enter that on the “NemFam id:” line. If you are interested in a particular Sanger EST contig, you can directly search on that. Doing so will return the NemFam id in which that contig is a member. The same for searching by EST read name. You will be returned the NemFam id for EST contig in which the EST read is a member.

The advanced search link leads to the form displayed on *slide 47*

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nematode_protein_families_advanced_search> (NemFam Advanced Search)

This allows the user to select a group of species to report on. You will be returned a list of NemFams that have at least 1 member from each selected species (i.e. selecting more and more species reduces the size of the return list).

Also shown on *slide 47* is the table that will be returned based on your requested species. The table lists all NemFams with at least one member from each selected species. This table also summarizes the available annotations for each returned NemFam group.

Clicking on one of the NemFam links takes you to a GBrowse view for that family. *Slide 48* gives an example of what you’ll find. The 2009 NemFam set was annotated for various structural features, all of which are displayed in their respective tracks.

Of note here is the fact that because NemFams are displayed against an artificially constructed spacer built independently for each family, normal GBrowse navigation requires you to know the exact NemFam id you want to jump to.

**Slide 49:**

This slide illustrates the interface used to access **NemaGene Transcript assembly** **data**

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=transcript_assembly_ftp> (NemaGene Transcript Assembly info. & download page)

**Slide 50:**

This figure illustrates the concept of isotigs & isogroups (all our 454 transcript assemblies were made using the Newbler transcript assembly, which provides isotig & isogroup information)

**Slide 51:**

This is our repository of **Codon Usage Tables**

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=codon_usage_tables> (Codon Usage Tables)

**Slide 52:**

This is our repository of **Sanger EST contigs**

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=cluster_data_ftp> (Sanger EST contig downloads)

**Slide 53:**

This slide describes the **Intestinal Transcriptome dataset**

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=intestinal_transcriptome> (Intestinal Transcriptome page)

**Slide 54:**

This slide displays our **Gene Expression Data** repository

URL:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=microarray> (Gene Expression Data repository)

**Slide 55:**

This slide is an overview of Nematode.net’s hosted datasets

**Slides 56-57:**

These slides describe the **NemaGene** repository. Here is a summary of direct URLs to resources supporting NemaGene.

URL:

**NemaGene Cluster Search**:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemagene_cluster_search>

**NemaGene Gene Table:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemagene_gene_tables>

**Codon Usage Tables:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=codon_usage_tables>

**AmiGO (GO Associations):**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=go_associations>

**NemaPath:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nemapath>

**NemaBLAST:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=nemablast>

**Intestinal Transcriptome:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=intestinal_transcriptome>

**Transcript Assemblies (454 assembly isotigs & isogroups):**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=transcript_assembly_ftp>

**EST Cluster Data (Sanger EST assembly contigs & clusters):**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=cluster_data_ftp>

**EST Read FTP (Sanger EST reads):**

\*note: Sanger EST reads are available for download via FTP. A description of how to access & navigate that FTP site is available towards the bottom of the page:

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=home&subnav_selection=data_ftp>

**NemaGene FAQ:**

http://nematode.net/NN3\_frontpage.cgi?navbar\_selection=nemagene&subnav\_selection=nemagene\_faq

**Slides 58-59:**

These slides describe the **HelmCoP** resource. Here is a summary of direct URLs to the main HelmCoP pages:

URL:

**HelmCoP front page:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=search_splash_page>

**HelmCoP Gene Search page:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=gene_search>

**HelmCoP Ortholog Search page:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=ortholog_search>

**HelmCoP BLAST service:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=search&subnav_selection=helmcop_blast>

**HelmCoP FAQ:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=helmcop_faq>

**Slides 60-61:**

These slides describe the **NemFam** resource. Here is a direct URL to the main NemFam page:

URL:

**NemFam main page:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=nematode_protein_families>

**Slides 62-63:**

These slides describe our **Gene Expression Data** repository. Here is a direct URL to the Gene Expression Data repository:

URL:

**Gene Expression Data:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=funcexp&subnav_selection=microarray>

**Slides 64-65:**

These slides describe our **Nematode Genome** **resources**. Here is a summary of direct URLs to these resources:

URL:

**NemaBrowse:**

http://nematode.net/NN3\_frontpage.cgi?navbar\_selection=nemagene&subnav\_selection=nemabrowser

**NemaSNP:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=nemagene&subnav_selection=pnema_snpdb>

**Slides 66-67:**

These slides describe our repository of **hosted publication data**. Here is a direct link to the Publication Data download area:

URL:

**Publication Data download:**

<http://nematode.net/NN3_frontpage.cgi?navbar_selection=home&subnav_selection=data_ftp>

**Slides 68-70:**

Upcoming projects information & Nematode.net credits.

