

Text comparison table: The left column contains text from the Gedela article cited below in blue. The right column contains references to works where portions of the Gedela text appeared previously, and below that, the actual text as it first appeared. Black text indicates matching text in both sources; gray text indicates text that occurred in only a single source.

None of the text included in the left column is attributed in the article, nor does any of Gedela's text included here appear within quotation marks in his book chapter.

Gedela, Srinubabu. (2011). Integration, warehousing, and analysis strategies of omics data . In B. Mayer (Ed.), <i>Bioinformatics for Omics data: Methods and protocols</i> (pp. 399-414). New York: Humana Press.	
Gedela, p. 407:	Borisjuk, L., Hajirezaei, M.-R., Klukas, C., Rolletschek, H., & Schreiber, F. (2005). Integrating data from biological experiments into metabolic networks with the DBE Information System . <i>In Silico Biology</i> , 5, 93-102. p. 97:
To characterize the responsible metabolic shift within seeds from sugars/starch into organic acids/amino acids/proteins, the metabolite pattern for glycolysis, citrate cycle as well as related sugars and free amino acids was analyzed.	To characterise the responsible metabolic shift within seeds from sugars/starch into organic acids/amino acids/proteins, the metabolite pattern for glycolysis, citrate cycle as well as related sugars and free amino acids was analysed.
Gedela, p. 408:	Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., & Kasprzyk, A. (2009). BioMart--biological queries made easy . <i>Bmc Genomics</i> , 10(22). p. 2:
BioMart (http://www.biomart.org) (9) is an open source data management system that comes with a range of query interfaces allowing the user to group and refine data based upon many different criteria.	BioMart is an open source data management system that comes with a range of query interfaces that allow users to group and refine data based upon many different criteria.
Gedela, p. 410:	Ibid, p. 5:

<p>The <i>Homo sapiens</i> genes data set is selected and filters of <i>ID List Limit</i> in the <i>GENE</i> section is chosen. Selecting the <i>Affy hg u95av2 ID(s)</i> option provides an upload option for Affymetrix probset IDs using the file Browse button, or alternatively by copy and paste of the data set into the text box. Data types include complementary DNA (cDNA), peptides, coding regions, untranslated regions (UTRs), and exons with additional upstream and downstream flanking regions. In order to identify upstream regulatory features in subsequent analysis, the 1 kb upstream flank sequence for each gene has to be selected (Fig 8).</p>	<p>The <i>Homo sapiens genes</i> dataset is selected, and filters selected by clicking on the Filter bar again but this time the <i>ID List Limit</i> filter in the <i>GENE</i> section is chosen. Choosing the <i>Affy hg u95av2 ID(s)</i> option allows the user to upload a file of experimentally relevant Affymetrix probeset IDs from this Genechip using the file Browse button or to enter IDs by cutting and pasting into the text box (we include some example IDs in Additional file <u>1</u>). The various sequence options can be seen by clicking on Sequences at the top of the page in the attributes section (Figure <u>2A</u>). These include cDNA (complementary DNA), peptides, coding regions, UTRs (untranslated regions), and exons with additional upstream and downstream flanking regions. In order to identify upstream regulatory features in subsequent analysis, the user would select 1000 bp of upstream flank sequence for each gene (Figure <u>2B</u>).</p>
<p>Gedela, p. 410:</p>	<p>Ibid, p. 9:</p>
<p>A number of external software packages have incorporated BioMart for enhancing querying capabilities, e.g. for using services as Galaxy, BioConductor, Taverna, or to add further annotation and visualization of results (e.g., Cytoscape, Http://www.cytoscape.org).</p>	<p>A number of external software packages have incorporated BioMart querying capabilities into their systems.</p>
<p>Gedela, p. 410:</p>	<p>Ibid, p. 9:</p>

<p>This integration has been made possible through MartServices. BioMart can be easily configured to become a DAS annotation server for viewing of data through various Distributed Annotation System (DAS) clients.</p>	<p>This integration has been made possible through MartServices. All the requests generated by these external packages run against the BioMart central portal. Using BioMart through these external packages expands the usefulness of both BioMart and these external tools. Therefore, brief descriptions and examples of this integrated usage are presented below. BioMart has also been improved by the incorporation of external software technologies.</p> <p>BioMart can be easily configured to become a DAS annotation server for viewing of data through various DAS clients.</p>
<p>Gedela, p. 411:</p>	<p>Wikipedia article for Grid resource allocation manager: [Note: The text in the article predates the publication of the Gedela article, and there is no evidence to indicate Gedela himself ever contributed to the Wikipedia article.]</p>
<p>3. Grid Resource Allocation Manager or Globus Resource Allocation Manager (GRAM) is a software component of the Globus Toolkit that can locate, submit, monitor, and cancel jobs on Grid computing resources. It provides reliable operation, stateful monitoring, credential management, and file staging. GRAM does not provide job scheduler functionality and is in fact just a front-end (or interoperability bridge) to the functionality provided by an external scheduler that does not natively support the Globus Web service protocols.</p>	<p>Globus resource allocation manager (or GRAM) is a software component of the Globus Toolkit that can locate, submit, monitor, and cancel jobs on Grid computing resources. It provides reliable operation, stateful monitoring, credential management, and file staging.^[1]</p> <p>GRAM does not provide job scheduler functionality and is in fact just a front-end (or interoperability bridge) to the functionality provided by an external scheduler that does not natively support the Globus web service protocols.^[2]</p>

Gedela, p. 413:	http://www.biotapestry.org/
BioTapestry is an interactive tool for building, visualizing, and stimulating genetic regulatory networks. The tool is also used for Interactive Web Models.	BioTapestry is an interactive tool for building, visualizing, and simulating genetic regulatory networks. Here are some uses of BioTapestry:

Similar figures: There are two figures that are somewhat identical in Gedela and Smedley, *et al.* I don't think that Gedela actually copied the figure; instead, it appears he may have gotten the idea of the figure from the earlier authors and created it himself using the BioMart software. This may be an example of copying someone else's ideas. [See next two pages].

Gedela, Figure 8 on p 411:

Integration, Warehousing, and Analysis Strategies of Omics Data 411

The image shows two screenshots of the bioMart web interface. The top screenshot displays the configuration page for selecting sequence attributes. On the left, under 'Filters for affymetrix data', the filter 'Affy hg u95av2 IDs (ID list specified)' is selected. Under 'Attributes', 'Ensembl Gene ID', 'Flank (Gene)', 'Upstream flank (1000)', 'Chromosome Name', 'Gene End (bp)', and 'Gene Start (bp)' are listed. The 'Sequences attribute' section is active, showing a diagram of a gene with exons and introns. Below this, 'Selected options' includes: 'Unspliced (Transcript)', 'Unspliced (Gene)', 'Flank (Transcript)', 'Flank (Gene)', 'Flank-coding region (Transcript)', 'Flank-coding region (Gene)', '5' UTR', '3' UTR', 'Exon sequences', 'cDNA sequences', 'Coding sequence', and 'Protein'. The 'Upstream flank' is set to 1000 bp. The 'Gene Information' section has 'Ensembl Gene ID', 'Chromosome Name', and 'Gene Start (bp)' selected. The bottom screenshot shows the 'Result of the above attributes and filters', displaying a FASTA format output with gene coordinates and sequence data.

Fig. 8. Example for sequence attributes, filters, and results after the selection of given options in the MART window.

Smedley, *et al.*, Figure 2 on p. 6:

BMC Genomics 2009, 10:22

<http://www.biomedcentral.com/1471-2164/10/22>

A

Dataset 698 / 36582 Genes
Homo sapiens genes (NCBI36)

Filters
Affy hg u95av2 ID(s): [ID-list specified]

Attributes
Flank (Gene)
Upstream flank [1000]
Ensembl Gene ID
Chromosome Name
Gene Start (bp)
Gene End (bp)

Dataset
[None Selected]

Please select columns to be included in the output and hit 'Results' when ready

Features
Structures
Variations

Homologs
Sequences

SEQUENCES:
Sequences (max 1)

Unspliced (Transcript)
Unspliced (Gene)
Flank (Transcript)
Flank (Gene)
Flank-coding region (Transcript)
Flank-coding region (Gene)

5' UTR
3' UTR
Exon sequences
cDNA sequences
Coding sequence
Protein

Upstream flank
 Upstream flank[1000]

Downstream flank
 Downstream flank

Header Information
Gene Attributes
 Ensembl Gene ID
 Description
 Associated Gene Name
 Associated Gene DB

Chromosome Name
 Gene Start (bp)
 Gene End (bp)

B

Dataset 698 / 36582 Genes
Homo sapiens genes (NCBI36)

Filters
Affy hg u95av2 ID(s): [ID-list specified]

Attributes
Flank (Gene)
Upstream flank [1000]
Ensembl Gene ID
Chromosome Name
Gene Start (bp)
Gene End (bp)

Dataset
[None Selected]

Export all results to FASTA Unique results only

Email notification to

View rows as FASTA Unique results only

```
>ENSG0000004487|1|23218533|23282771
TCACCCATAATCAAGATATAAAACATTCACATCACCCCGAGCAAGTCTTTCTTGTTCCTT
TGCAGTCAACCCCGACCCCACTCAGGCAAGGCTGATTGGAGCCTATTACTGGAGAT
TAGTTTACCCAGTCTTGAACCTCATGTAATGGAATAATTTTGGACAATTCACCTCG
AAGACACAGCTATTAAGAGCAATTTGGTAAATAGTACACGCATCCTAAAGTA
CACGACTGTCCAGTTTGGGTTAGCAAGGATTCCTCGAAAAATTAAGAATAAAC
CCPACTATCATATTAGAAGAATAACCTGGAGTAGAAGAGGTGTATGTGGATGGGG
ATGGGGACATGAAGATGGGGACATGGGGATGGGGACATGGGGATGGGAGAAGGC
CAAGACATGGAGTCGGAATGAGGTGGGGCCAGCCGAGAAAGCAATGAACAAAGCC
TTGGAGACAAGAGCAATTCCTCTGGACTCAAACCTGGATTTCGTAACCTGTCTCCATTT
TACTACCATCGAAACTGCTGTGCTCTCCAAGTTTTTGGCCAGTCCAAGCAGGCAATTA
TGAATAAGTAAGACCGCTACCAACATATTCAGGTTCCTCCCAATTAGACTGTAAGTT
CTTGAAGCTAGAACTCTATATCTAACCCTGTACCTAGTGGCAAGCAATTAATAA
TACTGGTTGAATTAACGAATAATGGGTGAGTGAATGAACGAACGCCAGCCTGCAAAACC
CAAAGTCCCTGGAGGAAATGGTCACTTCGGAGGTTTAGTCTGGCCGGAAGCCTAAGA
CCAGGACTGTGCCAGTCCCACTCAAACCCGGGGAGACGCTTAGGCAAGCTACACG
TTCTTTGCTGCGGTGCCACTTAGCCGCGAAGCCGCTCTATGGCTCGGGGGAGGGG
CGGGCTCGTGGSTGTCTCCGACCCCTTTTGTCCCGGGG
>ENSG0000000938|1|27811390|27834314
TGTGTTAGCCAGGACGCTGATCTCTCAACATCGTATCCACCCGCTCGGCCTCCCA
AAGTCTGGATACAGCGGTGAGCCCTGCACCCGGCCCTTTCTTCTTTTAAAAA
```

Figure 2

(A) Sequence output options and (B) FASTA output for all the genes found to be up-regulated in a microarray experiment using the Affymetrix HG-U95Av2 probeset. Here 1000 bp upstream of the first exon have been chosen along with the Ensembl Gene Id and the chromosomal position of the gene for the FASTA header.