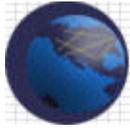




Meeting/Teleconference Name, Date

Gene CDE Focus Group Meeting Minutes

Date, Time & Location:	October 15, 2004 1:00-2:00 EDT																		
Attendees:	<table border="0"> <tr> <td>Juli</td> <td>Klemm</td> <td>3rd Millennium</td> </tr> <tr> <td>George</td> <td>Komatsoulis</td> <td>NCICB</td> </tr> <tr> <td>Rakesh</td> <td>Nagarajan</td> <td>Washington University</td> </tr> <tr> <td>Harold</td> <td>Riethman</td> <td>Wistar Institute</td> </tr> <tr> <td>Craig</td> <td>Street</td> <td>Penn</td> </tr> <tr> <td>Baris</td> <td>Suzek</td> <td>Georgetown University</td> </tr> </table>	Juli	Klemm	3rd Millennium	George	Komatsoulis	NCICB	Rakesh	Nagarajan	Washington University	Harold	Riethman	Wistar Institute	Craig	Street	Penn	Baris	Suzek	Georgetown University
Juli	Klemm	3rd Millennium																	
George	Komatsoulis	NCICB																	
Rakesh	Nagarajan	Washington University																	
Harold	Riethman	Wistar Institute																	
Craig	Street	Penn																	
Baris	Suzek	Georgetown University																	
Introduction:	Roll-call, review meeting goals (Craig Street)																		
caBIO Domain Objects	<p>caBIO Gene Domain Object (George Komatsoulis)</p> <p>caBIO models a gene to solve their particular use cases and not because it is necessarily the best model. They define genes from a data point of view using Unigene (one cluster = one gene). The gene object contains fields for title, LocusLink summary, OMIM, cluster ID, gene symbol, and hash table of database cross-references. All of the rest of the information is captured through association with other objects. GeneAlias, GeneHomolog, and Map Location are examples of other objects that are associated with a gene object. In general, the approach has been that if there can be more than one value, a separate object is used.</p>																		
Review Discussions:	<p>Gene CDE (Craig Street et al.)</p> <p>We are considering a long-term solution of a mapping service using a minimum set of identifiers. This will allow for a loose coupling of identifiers. There was some concern about having a transcript only database adding a new attribute for each entry which explicitly requires linking a transcript to the gene which encodes it. The mapping service would eliminate this requirement, as it could do the job of defining this relationship. There is also some concern about not drawing a distinction between a gene and its transcribed (RNA) and translated (Protein) products. The basic use case proposed is 1) Submit an ID; 2) Return the set of all possible identifiers for that object.</p> <p>Proposed definition: Gene=Composite set of identifiers linked by a mapping service.</p> <p>George suggested that it is common practice to assign a primary name for each object. A primary name (when defined) can enforce that for every gene there is a unique identifier. The HUGO Official Gene Name would fill this role. However, another 'key' would be needed for all genes objects, even when a HUGO name has not been assigned (the preferred name would not be the primary key). Note that HUGO names only apply to humans, not other organisms. An LSID is a possibility. This has been under consideration by the architecture group but not clear whether it has been fully sanctioned by caBIG. Juli to distribute Brian Gilman's LSID presentation.</p> <p>There was discussion around whether the ISO11179 metadata standard requires a preferred name for a data element – the discussion was inconclusive as to whether a preferred name is really metadata.</p>																		



Meeting/Teleconference Name, Date

	<p>The mapping service, if created, would return an object that contains all the accepted identifiers for that gene. In this way, if an individual submits a identifier that maps to more than one object, all are returned and the individual selects the appropriate one.</p> <p>GeneLnyx was provided as an example of an existing mapping services. It was suggested that this service could be put on the Grid.</p>			
Action Items:	Name Responsible	Action Item	Date Due	Notes
	Harold Riethman	Send potential list of gene and transcript identifiers.	Done	
	Baris Suzek	Send potential list of gene and transcript identifiers.	Done	
	Craig Street	E-mail developer project leads to determine what gene and protein identifiers tools are currently using.	Done	
	Craig Street/Rakesh Nagarajan	Revise white paper	Pending Decision of Group	
	George K.	Generate a PDF version of caBIO to share with the group	10/25/04	
	Juli Klemm	Send Craig and Rakesh Brian Gilman's LSID presentation	Done	