

BioPAX Work Group
November 3, 2004 Conference Call Minutes

Participants: Gary Bader, Michael Cary, Peter Karp, Joanne Luciano (see addendum), Suzanne Paley, Jeremy Zucker

Summary:

1. Recent ontology update (draft sent to biopax-discuss on Oct 18)
 - a. Summarized changes
 - i. Added disjointedness, sub-properties, restrictions
 - b. Release as revised Level 1?
 - i. No complaints or objections to releasing it
 1. Should CONTROLLED and CONTROLLER be PARTICIPANTS? Yes.
 - ii. We will release it as Level 1 version 1.1
 1. Suggested: If the modification does not affect data, increment version by +0.1; if the modification affects data, increment by +1.0
 2. Old versions should be made available online
2. Level 2 draft release
 - a. Gary and Mike working on draft, currently viewable on SourceForge CVS
 - b. Will likely be released later this week via biopax-discuss
 - c. Main features: hierarchical pathways, black-box pathways, molecular interactions
 - d. Review of new features
 - i. Hierarchical and black-box pathways (Jeremy)
 - ii. Evidence / support (Suzanne/Peter, et al.)
 1. Jeremy asked if computational evidence will be supported – it will, via external controlled vocabularies (e.g. BioCyc evidence codes)
 2. Currently no plans to incorporate such external ontologies in BioPAX, though it may behoove us to convert them (or request that they are converted) into OWL
 3. For the time being, we will continue to use instances of the externalControlledVocabulary class to create pointers to terms in external ontologies, rather than using URIs that point directly to those terms.
 - iii. We will solicit reviews of specific features based on interest, expertise
 - e. To do: Finish documentation (Mike), send out to biopax-discuss, solicit reviews, create examples
3. Data conversion
 - a. BioCyc progress
 - i. Next release: next week
 1. Will include a dump of the database in a single biopax file (one per organism)

2. Individual pathways?
 - a. Will not be available as separate biopax files for now
 - b. Desirable: a tool that would extract a single pathway from a large collection of pathways (in a BioPAX file)
 - c. Peter proposed a “generate biopax” button on pathway web page (long term)
 - i. Potential problem: unique IDs generated on-the-fly, so potential for rdf ID conflicts
 - ii. If the frame IDs were converted to rdf-compatible IDs in a consistent manner, that would solve the problem
 - iii. We may need to modify our best practice recommendations for IDs, depending on how semweb evolves
- b. Other datasets
 - i. KEGG
 1. Jeremy loaded KEGG into BioWarehouse
 2. Next: he will write a tool to generate BioPAX from a query to BioWarehouse
 3. Peter, Suzanne are updating their KEGG loader for BioWarehouse, so Jeremy will focus on the BioPAX generator tool
 - a. The only KEGG file Peter can find only has reactions with EC numbers – is there a file with reactions that lack EC numbers?
 - i. Jeremy will look into this
 - ii. Gary said they might be in the pathway files
 - b. Susumu Goto implied that we should translate KGML (KEGG Markup Language) into BioPAX
 - c. Jeremy tried using the KEGG API – works good but still need to parse the results (a flat file is returned)
 - ii. Reactome
 1. They are currently working on BioPAX export
 2. Gary did a mapping
 - a. Mike and Joanne discussed it with Peter D’Eustachio at SOFG (this was a follow-up from Joanne discussing it with Geeta Joshi-Tope at the DARPA BioCOMP PI meeting the week before).
 3. Jeremy suggested that they could publish their metadata in SBML as BioPAX (in addition to their making their data available as pure BioPAX)
 - a. Jeremy is working with the KEGG to SBML person
 - b. Jeremy will likely re-write the BioCyc to SBML converter to include annotations in BioPAX
 - iii. Other DBs
 1. WIT

- a. Joanne talked to Natalia
 - b. Natalia wants someone to work with them
 - c. Old WIT is gone, to be replaced with PUMA2 (see <http://compbio.mcs.anl.gov/puma2/cgi-bin/index.cgi>)
 - 2. Michael Jones (Novartis)
 - a. Offered his services to help with conversions
 - 3. Millennium folks (John Rees, Alan Ruttenburg) also expressed interest
- 4. Next meeting
 - a. Proposed date: January 21, 2005
 - i. One or two days? Depends on the agenda and how much we need to do, but one day is preferable
 - b. Location: tentative: NYC
 - i. Alternatives: Boston, SRI
 - ii. Depends on participants (e.g. Fukuda, Demir)
 - c. Agenda
 - i. Ratify Level 2, resolve any remaining issues
 - ii. Plan features of Level 3
 - 1. What do we want to include in Level 3
 - 2. Who will contribute proposals?
 - iii. Update roadmap
- 5. Future conference calls
 - a. First Wednesday of every month, 1 PM Eastern time
- 6. Recent meetings
 - a. Did not discuss due to lack of time, Jeremy and Mike will send meeting summaries to biopax-discuss
- 7. Miscellaneous topics
 - a. Linking pathway steps
 - i. Jeremy asked what links pathway steps together in the ontology. He said that currently you need to do a tortuous query to search for an intersection between participants between the two steps.
 - ii. Jeremy also mentioned that a developer that works on Haystack may have implemented a solution.
 - b. Physical interactions
 - i. Does this imply there will be non-physical interactions?
 - 1. Yes, genetic interactions (different from genetic-regulatory interactions)
 - 2. Genetic interactions will be a feature of level 3 or later
 - c. Level 2 release date?
 - i. Soon after next meeting (Jan 21) at the earliest
 - d. BIND support – they will support Level 2
 - e. Jeremy asked if we will incorporate any external CVs into BioPAX
 - i. Generally, no – better to simply point to external CVs rather than incorporating them
 - f. Future meetings: PSB

- i. Joanne sent in abstract to PSB; we need to determine who is going, who will represent BioPAX
- ii. Michelle Green from SRI is attending, Ken Fukuda may also; one of them may be able to represent BioPAX in some way (e.g. via fliers or presenting poster)

Addendum:

Joanne missed most of the call but expressed the following via email after reviewing the initial version of the minutes:

1. Future conference calls
 - a. Call times should be displayed more prominently on call announcements
 - b. We should more actively recruit conference call participants
 - c. Monthly calls may be too infrequent and may slow down progress
2. Data conversion effort
 - a. Should have a single central coordinator
 - b. Should have a subgroup with regular (probably weekly) meetings
 - i. Responsibilities would include development of libBioPAX
 - c. Notice should be sent to other pathway DBs telling them about BioPAX, asking about their level of interest
 - d. Reactome – what's the status of the mapping effort?
3. Miscellaneous
 - a. We could use a dedicated BioPAX project coordinator
 - b. The BioPAX webpage should contain direct and easy access to documentation and tutorials
 - c. In addition to Mike Jones, Vivien Bonazzi from Invitrogen, Jonathan Rees and Alan Ruttenberg are all standing by to help with conversion software, in addition to one or two others that may also be interested.