IODP Paleontology Meeting

Marriot Woodlands Courtyard, Houston, Texas, USA

September 29-30, 2006

Preliminary Agenda

Participants

Meeting Co-Chairs		
Soeding, Emanuel	IODP-MI	esoeding@iodp-mi-sapporo.org
Lazarus, David	MRC	david.lazarus@rz.hu-berlin.de
Participants		
Aita, Yoshiaki	MRC	aida@cc.utsunomiya-u.ac.jp
Alroy, John	PaleoDB	alroy@nceas.ucsb.edu
Blum, Peter	USIO	blum@iodp.tamu.edu
Huber, Robert	Stratigraphy.net	robert.huber@stratigraphy.net
Zarikian, Carlos	USIO	zarikian@iodp.tamu.edu
Graham, Colin	ESO	ccg@bgs.ac.uk
Huber, Brian	CHRONOS	huber.brian@nmnh.si.edu
Iwai, Masao	Kochi University	iwaim@cc.kochi-u.ac.jp
Van Couvering, John	Micropaleontology Press	vanc@micropress.org
Matsuda, Shigemi	CDEX	matsudas@jamstec.go.jp
Miville, Bernard	IODP-MI	bmiville@iodp-mi-sapporo.org
Pat Diver	CHRONOS	divdat@aol.com
Observers		
Aoike, Kan	CDEX	bluepond@jamstec.go.jp
Divins, David	JOI	ddivins@joiscience.org
Larsen, Hans Christian	IODP-MI	hclarsen@iodp-mi-sapporo.org
Gaillot, Philippe	CDEX	gaillotp@jamstec.go.jp
Yoshioka, Yuki	IODP-MI	yyoshioka@iodp-mi-sapporo.org
Mithal, Rakesh	USIO	mithal@iodp.tamu.edu
Goll, Bob	USIO	
Slone, Elizabeth	USIO	
Goll, Irina	USIO	
Houpt, David	USIO	houpt@iodp.tamu.edu
Mateo, Zenon	USIO	
Klaus, Ann	USIO	annklaus@iodp.tamu.edu
Fox, Jeff	USIO	fox@iodp.tamu.edu

Background

On a meeting in Kochi, Japan (January 2006), the Scientific Technology Panel (STP) recommended IODP-MI to coordinate the development of a Digital Taxa Dictionary (DTD) to improve IODP's own paleo data resolution and reproducibility, and ensure effective use of DSDP and ODP legacy sites. STP recommends that development of theses dictionaries should involve the Micropaleontological Reference Centers (MRCs), who could provide significant input to this process. IODPs data management coordination group at the same time discussed the use of taxa name lists (TNL), to enhance and streamline data capture through IODP onboard database systems. Both items, TNL and DTD can not be developed and maintained by the IODP alone. Hence there is a need to start a community effort to establish a system and protocols to compile, maintain and update such a list and dictionary.

Goals

The paleontology meeting is planned to be the start to develop an IODP strategy to approach some of the issues mentioned above.

Meeting goals regarding the TNL are to:
- decide on an TNL usable for IODP on all platforms with all databases
- define data formats
- organize the maintenance of the TNL
- promote this effort in the communities

Meeting goals regarding the DTD are:
- start the development process of a DTD
- define the community groups and projects that need to be involved
- define the technical infrastructure that will be needed
- define IODPs role in the development process

Background Material

Report of the SciMP (now STP) Paleontology working group

http://www.iodp.org/index.php?option=com_docman&task=doc_download&gid=131

Micropaleontology Reference Centers, Report to the IODP STP, *July, 2005*
In the STP agenda book, page 232pp.
http://www.iodp.org/index.php?option=com_docman&task=doc_download&gid=763

JANUS Paleo dictionaries
As Excel files: http://www.iodp.tamu.edu/database/paleo/
As taxon, datum, zone database for search:
http://www.iodp.tamu.edu/janusweb/paleo/paleo_dict.shtml

Main Agenda Items

Review existing paleontology projects:
- focus, goals
- community
- file formats
- data access

Agenda items related to the TNL:
- review currently available taxa name lists (i.e. the ODP taxa list)
- decide on an IODP name list that will be used on the platforms
- discuss the storage, format and accessibility
- discuss maintenance of the TNL

Agenda items for the DTD are:
- which community groups are interested in a DTD
- which community groups are already working on a DTD
- how can IODP contribute to the development of a DTD
- what technical infrastructure will be needed
- how can such a development/collaboration process be started

Pre-meeting action items

IODP-MI asks each participant to request input from their community about already existing databases. In particular information on their

- format
- storage
- access restrictions
- maintainers
- usage in the community

is essential. What other projects beside Paleodb and Chronos exist, and how can they be integrated.

Deliverables of the meeting

- an implementation plan to compile and maintain a TNL
- potential development plan for a DTD based on an integrated community effort.

Two weeks after the meeting an executive summary will be distributed to the participants. Six weeks after the meeting a draft report will be distributed to the participants.

Preliminary Agenda Schedule

Day 1: September 29, 2006

Time	Topic	Presenter
13h00	Introduction – Logistics – Participants	Soeding, Lazarus
13h30	Introduction to the goals of the meeting	Soeding, Lazarus
	IODPs current role in paleontology projects	
13h45	USIO current procedures and community links	Zarikian
14h00	ESO current procedures and community links	Graham
14h15	CDEX current procedures and community links	Aita
	Relevant Paleo Projects, Groups and Initiatives	
14h30	MRCs	Lazarus
14h50	Micropaleontology Press	Van Couvering
15h10	PaleoDB	Alroy
15h30	Break	
16h00	Chronos	B. Huber
16h20	GBIF/TDWG, GUID, ICZN Name Bank	R. Huber
	Phase I: IODP Immediate Needs	
17h00	Begin Discussion on TNL (need, existing ones)	all
17h30	End of meeting	

Summary:
The goal of this afternoon is, to compile all information about the program, IOs and relevant projects, so everybody can enter the discussion in the evening with the same background information. Therefore **Soeding/Lazarus** will give a brief introduction on what the goals and purpose of this meeting are, about immediate IODP-MI needs, and future projects involving IODP and community interaction.

After the introduction one representative for **each IO** will give a brief presentation on current procedures capturing paleontologic data, what dictionaries are used, what constraints applied. It should also be addressed what results possible community surveys had, and if there were any community requests for certain procedures and changes to existing data capturing systems. In addition any links to other projects or communities should be highlighted, where does IODP already influence paleontologic research, where do collaborations exist with database projects.

Subsequently **paleontologic databases, projects and initiatives** are asked to explain what their focus is, and where this focus overlaps with IODP science. It should also be addressed how the problem maintaining consistent taxonomic catalogues and databases was solved within these projects.

Phase I topics: immediate IODP needs: Taxa name list (TNL)

TNL for IODP onboard use. Discussion how to compile, maintain, store, collaborations. Suggestions/Actions on these questions:

1. Is the ODP list suitable to start with? If not, what is needed? Are there other existing TNLs?
2. Should this list be stored in a central place, keeping it accessible for scientists? Should people be able to change it or comment on it?
3. Is this list useful for the IOs to verify taxa names in the database.
4. Will a list like this be accepted by the scientific community? What problems can be foreseen, how can paleontologists be convinced to use it? Which are the major problems learned from ODPs JANUS?
5. Through which process can the correctness and accuracy of the TNL be maintained and how can new names be added? How do other projects handle this?
6. In which common data format should this list should be kept in (i.e. Excel, XML) to be easily exchangeable with other projects and/or scientists?

IODP seeks broadest possible consent on all these items with the different paleontology projects working on similar tasks.

Day 2: September 30, 2006

Time	Topic	Presenter
9h00	Introduction (if needed)	Soeding
9h10	Continue Discussion on TNL (how to setup, how to maintain)	all
10h30	Break	
11h00	Summarize Discussion on TNL	Soeding
	Phase II: IODP Future Needs, DTDs	
11h15	Summary on DTDs (why, how, who)	Lazarus
11h45	Begin Discussion on DTD (which exist, formats)	all
12h30	Lunch	
13h15	Continue Discussion on DTD (how to develop how to maintain)	all
15h30	Break	
16h00	Summarize Discussion on DTD	Soeding
16h20	Discussion: How do we proceed?	all
17h30	End of meeting	

Phase II topics: possible future IODP needs (DTD)

1. Presentation on Digital Taxonomic Dictionary: which ones do already exist, what format, who developed, how maintained, how distributed?
2. Discussion about the DTDs: Who needs a DTD? Why? Which expertise is needed to compile it? Which communities are best suited to provide this expertise? Who should take the lead?
3. What can IODP contribute to the development of a DTD and how can external groups be involved in the creation and maintenance of a global taxonomy dictionary/atlas?
4. What technical infrastructure would be needed for setting up a web based searchable atlas that can be maintained by the community. How can this data be exported into other project databases.
5. How can we start such a development/collaboration process?

Report of the IODP Paleontology Meeting

Marriot Woodlands Courtyard, Houston, Texas, USA

September 29-30, 2006

Updated July. 18, 2007

Participants

Meeting Co-Chairs		
Soeding, Emanuel	IODP-MI	esoeding@iodp-mi-sapporo.org
Lazarus, David	MRC	david.lazarus@rz.hu-berlin.de
Participants		
Aita, Yoshiaki	MRC	aida@cc.utsunomiya-u.ac.jp
Alroy, John	PaleoDB	alroy@nceas.ucsb.edu
Blum, Peter	USIO	blum@iodp.tamu.edu
Huber, Robert	Stratigraphy.net	robert.huber@stratigraphy.net
Zarikian, Carlos	USIO	zarikian@iodp.tamu.edu
Graham, Colin	ESO	ccg@bgs.ac.uk
Huber, Brian	CHRONOS	huber.brian@nmnh.si.edu
Iwai, Masao	Kochi University	iwaim@cc.kochi-u.ac.jp
Van Couvering, John	Micropaleontology Press	vanc@micropress.org
Matsuda, Shigemi	CDEX	matsudas@jamstec.go.jp
Miville, Bernard	IODP-MI	bmiville@iodp-mi-sapporo.org
Pat Diver	CHRONOS	divdat@aol.com
Observers		
Aoike, Kan	CDEX	bluepond@jamstec.go.jp
Divins, David	JOI	ddivins@joiscience.org
Larsen, Hans Christian	IODP-MI	hclarsen@iodp-mi-sapporo.org
Gaillot, Philippe	CDEX	gaillotp@jamstec.go.jp
Yoshioka, Yuki	IODP-MI	yyoshioka@iodp-mi-sapporo.org
Mithal, Rakesh	USIO	mithal@iodp.tamu.edu
Goll, Bob	USIO	goll@iodp.tamu.edu
Slone, Elizabeth	USIO	slone@iodp.tamu.edu
Goll, Irina	USIO	goll_i@iodp.tamu.edu
Houpt, David	USIO	houpt@iodp.tamu.edu
Mateo, Zenon	USIO	mateo@iodp.tamu.edu
Klaus, Ann	USIO	annklaus@iodp.tamu.edu

2. Executive Summary

IODP-MI organized a paleontology meeting in Houston, Texas, to tackle some immediate and long term IODP needs in Paleontology and start a dialogue with different community projects/initiatives to explore common solutions. This was in part in response to STP recommendations and to insure adequate paleontology data quality for use in various IODP applications. This meeting concentrated on primary paleontology data recording (taxonomic names and their documentation) rather than the application of these data , e.g. biostratigraphy, paleoceanography. IODP is making use of the expertise and links the MRCs have to the paleontologic community as well as to IODP. IODP paleontologic taxonomic work needs to be linked to similar community interests/projects. The meeting recognized the importance of IODP as one of the major programs contributing to micropaleontology, hence networking of a community effort is appropriate.

A step-wise development of 'Taxonomic Reference Standards' was determined to be the most effective approach, based on levels of information:

Level 1 - Taxonomic Name List "TNL" - create a comprehensive but standardised list of taxa names. This is needed for any meaningful entry or subsequent use of paleontology data in IODP.
Level 2 - "enhanced TNL" - Add synonyms. This is needed to retrieve data in a useful form.
Level 3 - " Digital Taxonomic Dictionary (DTD) - Add definitions and images - Needed for adequate Quality Assurance in data entry.
("Level 4" includes data, e.g ecology, only needed for post cruise science goals).

We recommend the following actions relevant to IODP, but designed to benefit the whole community.
1. Create TNL for IODP-wide immediate use (merge existing lists, ODP and other e.g. Neptune at Chronos - IODP-MI coordinated).
2. (Community) Web Portal to access existing (external and possibly proprietary) data at various Levels, particularly existing Level 3 (DTD) data resources.
3. Set up a paleontology coordination group (PCG) to advise/assist in these and following tasks.
4. IODP establish a policy for adequate referencing of species names in IODP publications thereby linking all paleontology name uses to the TNL.
5. Organize maintenance/updates of the TNL through taxonomic editors and/or a standing names board.
6. Identify regional and spatial gaps in coverage of external taxa databases and promote work on these in the community (possibly using MRC resources, and preference given to gaps related to existing proposals).
7. In addition to cruise gathered paleontology data, IODP should also establish a mechanism to capture post-cruise data.

The meeting participants also recommend to IODP-MI to coordinate storage and dissemination of other biostratigraphic data (e.g. occurrences of taxa in samples) among IOs and within the community.

3. Minutes

In January 2006 the IODP Scientific Technology Panel (STP) recommended IODP-MI to coordinate the development of a Digital Taxa Dictionary (DTD) to improve the accuracy and reproducibility of IODP paleontologic data, and ensure effective use of DSDP and ODP legacy sites. STP mentioned that the development of theses dictionaries should involve the Micropaleontological Reference Centers (MRCs), who could provide significant input to this process. IODP's data management coordination group at the same time discussed the use of taxa name lists (TNL), to provide coherent paleontology data capture throughout IODP onboard database systems. TNL and DTD can not be developed and maintained by the IODP alone. Hence IODP-MI thought that there is a need to start a community effort to establish a system and protocols to compile and maintain such a list and dictionary. To initiate this effort an IODP Paleontology Meeting was held in Houston (Texas), which brought together representatives of various paleontological community databases, the MRCs, IODP implementing organizations, data management experts and IODP-MI staff, to discuss how this can be accomplished.

Meeting goals related to the TNL were to:
- decide on a TNL usable for IODP on all platforms with all databases
- define data formats
- organize the maintenance of the TNL
- promote this effort in the communities

Meeting goals related to the DTD were to:
- start the development process of a DTD
- define the community groups and projects that need to be involved
- define the technical infrastructure that will be needed
- define IODPs role in the development process

After an introduction to the goals and scope of the meeting by Soeding and Lazarus, participants presented and discussed various paleontology related projects, initiatives and developments, to form a common basis of information (see Appendix A). Presentations started from the IODP IO's reporting on their paleontology data handling, followed by presentations on other existing databases and programs, among them PBDB, Chronos and Micropaleo Press. Finally current progress on global taxa ID standardization was shown, initiated by the International Union of Biological Sciences Working Group on Taxonomic Databases (TDWG).

During discussions it became clear that the terms taxa name list (TNL) and digital taxonomic dictionary (DTD) were variants of the same basic concept and needed a common terminology definition. The group came up with a definition defining the TNL as a simple starting point of taxa names, which, by adding more kinds of data, can be evolved into a DTD:

Terminology TNL, DTD
Taxon Name Lists (TNL) and Digital Taxonomic Dictionaries (DTD) are Taxonomic Reference Standards in any format, containing taxonomic (and for IODP, particularly species) information.

Taxonomic Reference Standards can come in several levels:
#1. TNL
As a list of fossils containing the following information
• Genus, subgenus, species, sub species, variety
• Reference (Author, Year)

#2. extended TNL
like #1. +
• synonyms

#3. DTD
like #2. +
• Original description, remarks, repository (curatorial Information. Type locality, type level)
• Differential diagnoses (Comparative)
• Pictures
• stratigraphic information
• spatial distribution

#4. extended DTD
like #3. +
• More advanced information like
• ecological information
• morphologic descriptors
• and others....

The group therefore concluded that TNL and DTD are members of the same data development effort, so it would be appropriate to search for a common stepwise solution, building a TNL (#1) first from existing data sources, continuing stepwise through this process until a DTD stage (#3) can be reached. It should be possible to capture necessary data for the extended DTD (#4) as well as it appears.

The group agreed that the problem is to find a general solution for IODP, which makes use of already existing community resources, but also fills in gaps necessary to keep IODP operations on a high quality level. As a solution for this it was proposed:
- to set up a Taxa-Wiki or Portal as an access point to collect information about species.
- to attach metadata to the information (so it can meaningfully be searched)
- to make use of universal tags if available such as the proposed LSID (Life science ID)

- collaborate and query information in other databases
and
- install an IODP Paleontology server, where other projects could store their data as well.
or
- add links to these to the Wiki data (preferably through LSID)
and
- make everything exportable in various formats, so it can be used by other databases

To estimate whether this idea was worthwhile to follow up on, the group then compiled a list of existing databases, publications and projects, already containing paleontology data, and appended which level (TNL to DTD) this data had already reached (Appendix B). A goal would be to contact these databases and compile an initial TNL from the data present. The group agreed that a combination of Neptune (at Chronos) name data and ODP Taxa Name data would make a good start for a simple TNL, which would be worthwhile for shipboard and database use.

Action Item 1: IODP-MI should contact maintainers of these databases, investigate status and negotiate the usage of the data. (Timeline: by January 31, 2007)

The immediate IODP Goal/Need is a development plan for Taxonomic Reference Standards based on staged taxonomic information approach e.g. levels:
- 1 compile comprehensive list of names & their authors -needed for anything
- 2 add Synonyms -needed to retrieve data meaningfully
- 3 add descriptions and images - needed for adequate Quality Assurance in data entry
- (4 data for post cruise science goals)

We recommend the following procedures to accomplish this. These are all relevant to IODP, but designed to benefit and enhance cooperation with the whole science community.

A. Create a TNL for IODP immediate use by merging existing lists, ODP and others. This should be IODP-MI coordinated:

Action Item 2: IODP-MI should collect this external data, starting with the ODP and the Neptune TNL level 1/2 data, combine the lists, and coordinate with the MRC's to validate the result. Then pass it back to the IOs for implementation. Details of implementation should be coordinated by the IODP data management coordination group (DMCG). (Timeline: by February 28, 2007).

B. (Community) Web Portal to access (external and possibly proprietary) data at various levels

Action Item 3: IODP-MI shall devise an implementation plan how to organize setup of such a portal. (Timeline: by February 28, 2007).

C. Form a paleontology coordination group (PCG)

Action Item 4: IODP-MI shall create a Paleontology coordination Group (PCG), which should liaise with the MRCs and broader paleontology community, and advise on issues related to the paleontology portal or other paleontology or biostratigraphy related issues. (Timeline: membership by January 31, 2007, first meeting in second quarter 2007).

D. IODP should establish a policy for adequate referencing of species names in IODP publications (ensure that name usage can be linked back to the new TNL):

Action Item 5: IODP-MI should establish a policy for adequate referencing of species names in IODP publications, to make sure acquired paleontology data can be checked against the TNL and future DTD and thus meets necessary quality standards. (Timeline first quarter of 2007)

E. Organize maintenance/updates of the Taxonomic Reference Standards through taxonomic editors and/or via a standing 'taxonomic names board' (maybe with incentives)

Action Item 6: With advice from the PCG, IODP-MI should establish the necessary organization/mechanisms to ensure that Taxonomic Reference Standards used by the program are regularly updated to reflect new scientific knowledge. (Timeline: N/A, after first PCG meeting)

F. Identify regional and spatial gaps in coverage of taxa databases and promote work on these to the community (preference given to gaps related to existing proposals). No action item as yet -Details to be worked out by PCG.

G. In addition to cruise gathered paleontology data, IODP should also try to capture post-cruise data. No action item her yet - details need to be worked out by the PCG.

The meeting participants also noted to IODP-MI that other paleontology related issues, such as the coordinated storage, dissemination and integration of biostratigraphic data among IOs and the community will still need to be

addressed at a later date. A detailed discussion of such issues is contained in the STP Paleontology Working Group report from spring 2004 (see Appendix C).

Appendices

Appendix A
List of Presentations

USIO current procedures and community links	Carlos Alvarez Zarikian
ESO current procedures and community links	Colin Graham
Stratigraphy at 902-C9001 Chikyu '06Sep	Kan Aioke
Paleontological databases developed in Japan	Yoshiaki Aita
The MRCs – a few quick comments	David Lazarus
Micropaleontology Press and other databases	Van Couvering
Paleobiology DB - Dynamic Online Taxonomy System	John Alroy
Planktonic Foram Digital Taxonomic Atlas	Brian Huber
CHRONOS, Foram Catalogs and Neptune	Patrick Diver
Identifiers for the Life Sciences	Robert Huber

Appendix B
List and status of existing databases
See compiled_taxon_databases.xls
UPDATE: This is a constantly updated document, as we become aware of these databases, so please request the latest list from IODP-MI, esoeding@iodp-mi-sapporo.org.

Appendix C
STP Paleontology Working Group report from spring 2004
Paleo_MRC_WG.pdf
(see document online at
http://www.iodp.org/index.php?option=com_docman&task=doc_download&gid=131)