

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

A. Executive Summary	p. 1
B. UniEuk Vision & Principles	p. 5
C. Next Steps	p. 6
D. UniEuk Timeline	p. 11
E. Meeting Notes	p. 12
F. Meeting Agenda	p. 22
G. Participants List	p. 24
H. UniEuk Committees	p. 25

A. Executive Summary

The UniEuk project was officially launched in May 2016, with seed funding from the Gordon & Betty Moore Foundation (GBMF) and the International Society of Protistologists (ISOP). A kick-off and first implementation meeting was organized at the Museum National d'Histoire Naturelle (MNHN), Jardin des Plantes, Paris, May 3-5 2016, with the main goal to make consensus-driven and practical decisions about the implementation of UniEuk. All together (sections G and F), we determined how and when the key tools, modules, and stages of UniEuk should take place, and refined the planned community resources and outcomes. A list of goals and deliverables for the first 6-12 months was set up. The meeting was also an opportunity to reconfirm a common UniEuk vision, discuss the project's timeline and future meetings dates, and refine our outreach strategy.



The 32 participants to the UniEuk Kickoff Meeting, Paris, MNHN, May 2016.

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

1. UniEuk Principles

Throughout the meeting, spirited discussions drew upon participants' different perspectives and resulted in a general consensus on what the UniEuk project is and is not. We also returned the project's primary objectives, and a guiding principle was formulated and agreed upon by all participants (Section B).

2. UniEuk Committees

- The roles of the Steering Committee (SC), Advisory Council (AC), and Scientific & Technical Advisory Board (STAB), as defined in the project proposal, were reiterated (Section F).
- Based on group consensus among all meeting participants, including all SC members, the following changes to the UniEuk committee structure were made: (i) David Caron was selected to serve as chair of the AC; (ii) Juliet Brodie (current president of the International Phycological Society) and Stefan Geisen (leader of the Soil Protists Initiative) are invited to join the SC; (iii) Alexandros Stamatakis (Heidelberg Institute for Theoretical Studies, Germany; sco.h-its.org/exelixis/) is invited to join the STAB.
- Given the broader opportunity of UniEuk to ultimately include all eukaryotes, the AC should be expanded and include additional scientific authorities in plant, animal, and fungal (PAF) taxonomy/biodiversity, and databasing in Life Science (DNA and images) (Section C, WG8). Focus for now will remain on protists.



The UniEuk Steering Committee (section H), with, from left to right: Eunsoo KIM, Alastair SIMPSON, Stefan GEISEN, Sina ADL, Virginia EDGCOMB, Colomán DE VARGAS, Pelin YILMAZ, Guy COCHRANE, Javier DEL CAMPO, Frank Oliver GLÖCKNER

3. UniEuk Workflow and Taxonomic Framework

The meeting was an opportunity to reach a consensus agreement on the UniEuk workflow, its inputs and outputs, and to refine some general principles about the UniEuk flexible, adaptive, online taxonomic framework.

- The UniEuk taxonomic framework includes both described (sequenced or not-sequenced) and undescribed (environmental OTUs) taxa, as per the objectives of the project.

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- ^{18S}EukRef outputs (curated and annotated 18S rDNA alignments and phylogenetic trees) will form the backbone of UniEuk, but inputs to the UniEuk workflow will also include High-Throughput-Environmental-Sequence (HTES) data, single-gene phylogenies based on other genetic markers, phylogenomic data, and available information on described taxa and their morphology from the literature and image databases, etc.
- It is important to make the distinction between phylogenetic trees resulting from ^{18S}EukRef pipelines & Evolutionary Placement Algorithms (EPA), and the UniEuk taxonomic framework, which is a composite phylo-taxonomic map summarizing genetic and morphology-based information on protist biodiversity, taxonomy, and phylogeny. We decided to call it 'UniEuk taxonomic map (taxo-map)' to avoid confusion.
- Community feedback on the UniEuk taxo-map will happen via a navigable tree editable by lead-experts and linked to a wiki system allowing comments, links to references and image data, etc.
- The UniEuk taxo-map should use unlimited numbers of ranks, so we will have to find solutions to ensure compatibility with end-users needing a fixed number of ranks to operate.
- UniEuk's main output is the adaptive taxonomic framework editable by the world community of eukaryote experts and implemented into the International Nucleotide Sequence Database Collaboration (INSDC), beginning with EBI. The UniEuk e-taxomap thus allows continuous phylo-taxonomic revisions -matching the rapid evolution of current knowledge on eukaryotic biodiversity and phylogeny -, while being regularly frozen via versioning implemented at EBI to serve as universal reference.

4. Implementation Groups & Next Steps

All meeting participants agreed upon the creation of a few key, small, and efficient Implementation Groups (IG) that will carry out the practical tasks identified as next steps during the discussion sessions (section F). Seven IGs are proposed for now and presented in Section C; more can be created over the course of the project as appropriate.

In each group we suggested possible participants and highlighted a prospective group leader (in bold) who will make sure that all tasks and deliverables are achieved on time. People can be part of more than one group. For all identified tasks, prospective deadlines are also proposed, with possible tuning by the group members if necessary. Each IG should include at least one AC and/or STAB member who accepts to act as a reporter and will present an update of the IG's advances at the next UniEuk committee meeting (see "UniEuk Timeline" below).

Please check Section C to see whether you have been suggested as a member of one or more IGs, and for detailed information about the specific tasks of each group.

Each Implementation Group will be responsible for:

- Ensuring the appropriate people (including any relevant ones outside of UniEuk) have been invited to join the group (suggested limit of max. 8-10 people per group)
- Making sure people participating in the group commit to help achieve the deliverables on time
- Addressing all tasks and pending questions identified for their group
- Organising virtual meetings and/or (sub)group workshops (see "UniEuk Timeline" below; it is up to each group to decide if and when it is useful for them to meet in person)
- Reporting to all UniEuk committees and the other Implementation Groups

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

5. UniEuk Timeline

Here are the main updates / decisions taken (see the infographic in Section D):

- Pelin Yilmaz successfully submitted a proposal for a UniEuk roundtable at the 16th ISME meeting in Montréal, Canada, Aug 2016 (<http://www.isme-microbes.org/isme16>)
- Instead of another large implementation meeting during year 1, each of the Implementation Groups (Section C) will meet virtually and/or (if they decide it is more useful) in person over the next 6-12 months to address the specific tasks they are responsible for; as many such small workshops can be organised as necessary, in the limit of our current budget. Meeting planning must be coordinated with Colomban and Cédric, who are managing the central project funds.
- A second ^{18S}EukRef workshop will take place in Barcelona, during the week Nov 28-Dec 2, 2016. The organismal focus is currently expected to be Amoebozoa and Rhizaria, and perhaps other groups.
- At the same time (Nov 29-30, 2016), a UniEuk committees meeting will take place to assess the project's progress after 6 months; a main goal of this meeting will be for reporters of the Implementation Groups (see point 2 above) to present an update of the IGs' advances.
- The ICOP 15 meeting in Prague, Czech Republic, Jul/Aug 2017 (<http://www.icop2017.org>) will be the ideal venue to unite all UniEuk committee members, present to the community a first version of the UniEuk taxo-map online (with a couple of case-study), and advertise widely our first results.
- Possibility to submit a workshop proposal for the 11th International Conf. on Dinoflagellates, Bordeaux, France, Jul 2017 – opportunity to hold a ^{18S}EukRef workshop on dinoflagellates (?).

6. UniEuk Outreach

The meeting was also an opportunity to refine our outreach strategy.

- The next wave of outreach to potential stakeholders / end-users must be goal-oriented; prior to contacting them, we must determine clearly and be ready to explain: (i) What we want to get from them; (ii) What benefits they will get from liaising with UniEuk.
- We will prepare appropriate communication material in advance of outreach activity:
 - UniEuk website and short white-paper (1-2 pages)
 - Different type letters to contact different kinds of stakeholders.
 - Set of UniEuk slides that we can all use to present / advertise the project at meetings, etc.
- We will prepare use case scenarios that will become a communication device to clarify the UniEuk mission, help explain how UniEuk is integrative, what it is, what it is not
 - Everyone internally will have examples of how the UniEuk pipelines and workflow work
 - External stakeholders will understand the UniEuk mission, structure, and implementation
- We should create and maintain a UniEuk wish-list of resources, actions, side projects, etc. that we want to encourage in parallel to core UniEuk deliverables over the next 3 years.

7. Other Decisions Taken

The meeting's discussion sessions were organised around six main topics (Section F). Section E provides full details of all discussions, including for each topic lists of updates, decisions taken, and pending questions/next steps identified.

B. UniEuk Vision & Principles

Vision **UniEuk: Naming Life's Majority**

Universal taxonomic frameworks have been critical tools to structure the fields of botany, zoology, mycology, and bacteriology as well as their large research communities. Animals, plants and fungi have relatively solid, stable morpho-taxonomies built over the last 3 centuries, while bacteria and viruses have been classified for the last 3 decades under coherent molecular taxonomic frameworks. By contrast, no such common language exists yet for microbial eukaryotes (protists), even though environmental *-omics* surveys are showing today that protists make up most of the organismal and genetic complexity of our planet's ecosystems! In fact, the extreme morpho-functional complexity of protists has historically divided the relatively small research community into sub-communities (protozoology vs. phycology, aquatic vs. terrestrial protistology, fossil vs. modern protistology, etc.) speaking different languages. With the current deluge of eukaryotic meta-omics data, we urgently need to build up a universal eukaryotic taxonomy bridging the protist *-omics* age to the fragile, centuries-old body of classical knowledge that has effectively linked protist taxa to morphological, physiological, and ecological information. We thus propose UniEuk, a new online, integrative (morpho-genetic) and adaptive taxonomic system (e-TaxoMap), editable by the world community of eukaryote experts and implemented into the International Nucleotide Sequence Database Collaboration (INSDC), beginning with EBI. The UniEuk e-TaxoMap allows continuous phylo-taxonomic revisions - matching the rapid evolution of current knowledge on eukaryotic biodiversity and phylogeny -, while being regularly frozen via versioning implemented at EBI. In the long run, it should inform major scientific databases (DNA, images, etc.), and provide a universal taxonomic framework for eukaryotic biology, evolution, and ecology, both in research and education.

Guiding Principles

We are an open and inclusive, expert-driven initiative aimed at developing a universal and evolvable taxonomic framework for eukaryotes, with a focus on protists. We apply a transparent decision-making process that requires compromise and consensus-building. Our efforts incorporate DNA sequence phylogeny, morphological information (images), and other relevant data to build a robust, yet flexible and integrative working taxonomy of eukaryotes for the benefit of the broader community.

Near-term (3-year) opportunity

All protists can be effectively identified, classified, and compared using a universal integrative taxonomy that describes their evolutionary history and relationships. The UniEuk taxonomic framework is designed to integrate plants, animals and fungi when the time comes.

Primary objectives

1. Build consensus on an open, editable and evolvable, eukaryotic taxonomic framework via international community workshops and an interactive online system.
2. Optimize the shared taxonomic map by taking into account relevant sequence comparisons and morphological information for expert-defined taxonomic groups
3. Use the resulting taxonomy to inform major scientific databases (of genetic markers, images, etc.) and make it available to INSDC and other stakeholders, including basic educational textbooks, schools, and museums.

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

What UniEuk is:

- First priority = focus on overall eukaryotic taxonomy (“do one thing and do it well”)
- A high-priority effort, very pragmatic, but not necessarily all encompassing (high quality over comprehensiveness); a flexible and adaptive system
- An effort to organize eukaryotic names and taxonomy from the already-existing collective knowledge in the community
- An effort to give preliminary, standardized names (codes) to the majority of eukaryotic biodiversity which is still nameless (purely environmental sequences)
- A dynamic, consensus taxonomic framework representation that evolves in response to developments and community input (in the longer-term: will gradually include other groups of organisms to improve the system)
- A common reference for interpreting all sorts of genetic and morphological eukaryotic data
- A single, integrative (different types of information) framework to facilitate studies of ecology and environments, which include the biocomplexity of eukaryotic life.

What UniEuk is not:

- An effort to create analytical platforms for downstream analysis (but some may be by-products of the project)
- Creating new reference genetic databases (but will actively link to existing and future ones *)
- An image database (not asking the world for all of their images)
- A forum or vehicle where people can hash out very specific issues of taxonomy (“if we do that, it will never reach the end goal”)
- Perfect: at the cost of being pragmatic, there are going to be imperfections; we have to accept that and move on

* The ^{18S}EukRef component of the broader UniEuk effort will generate new reference material (alignments, trees, etc.) shared with existing genetic databases, while UniEuk overall is a taxonomic map – not a database – because it will incorporate multiple gene markers and morphological information.

C. Next Steps

As explained in Section A, all meeting participants agreed upon the creation of a few critical Implementation Groups that will carry out the practical tasks that were identified as next steps during the discussion sessions. Additional tasks that do not necessitate a dedicated IG but must be completed during the project's first 6 months are also listed below.

Note: The UniEuk committees (SC, AC and STAB) can be seen as Working Groups themselves; they can be attributed specific tasks (in particular, fundraising in case of the SC) and will have dedicated mailing list addresses (sc@unieuk.org, ac@unieuk.org, stab@unieuk.org).

Implementation Group 1 - ^{18S}EukRef pipelines / email: eukref-pipes@unieuk.org

Suggested Participants:

Javier del Campo, Laura Wegener-Parfrey, Stéphane Audic, Christian Quast, Frédéric Mahé, Lucas Czech, Oleksii Kozlov, Daniel Richter, Pelin Yilmaz, Colomban de Vargas

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

Specific tasks already identified:

- Check the ^{18S}EukRef pipelines and send comments
Frédéric Mahé, Christian Quast, Lucas Czech, Oleksii Kozlov
Proposed deadline = 15 June 2016
- Implement ^{18S}EukRef pipelines in the Galaxy instance at SBR
Stéphane Audic
Proposed deadline = end of September 2016
- Liaising with the UniEuk implementers to share existing and future ^{18S}EukRef outputs
Javier del Campo
Proposed deadline = end of September 2016

Pending questions to address:

- Replacement of uchime by vsearch?
- ^{18S}EukRef being clade-based, how to identify existing INSDC 18S sequences outside of known clades so they can also be taxonomically annotated?

Implementation Group 2 - Swarm Bank / email: swarm-bank@unieuk.org

Suggested Participants: **Colomban de Vargas**, Frédéric Mahé, Cédric Berney, Stéphane Audic, Lucas Czech, Oleksii Kozlov, Pelin Yilmaz

Specific tasks already identified:

- Empirical tests of merging different V4 rDNA HTES datasets into a single set of swarms
- Identify existing V4 datasets and contact relevant people to access them if necessary
Frédéric Mahé, Colomban de Vargas, Stéphane Audic, Cédric Berney
Proposed deadline = end of September 2016
- Define an EPA-based strategy to plug the novel protistan HTES data into a coherent phylogenetic tree - before their integration into the UniEuk taxo-map.

Pending questions to address:

- Online public interface, exploration and visualization of the Swarm Bank
- What do we want users to be able to do? (useful output would be to allow comparisons between primers / extraction methods used / etc.)
- How to provide the information associated with the raw data? (origin etc.)

Implementation Group 3 - Navigable Taxonomic Map, Wiki System, & Datamodel

/ email: taxo-map@unieuk.org

Suggested Participants:

Christian Quast, Bremen implementer, Cédric Berney, Stéphane Audic, Alastair Simpson, Philippe Deschamps, Colomban de Vargas, Pelin Yilmaz, Frank Oliver Glöckner.

Specific tasks already identified:

- Create a mock-up interface (navigable map & wiki) and have it appraised by 3-4 testers
Christian Quast / testers: Cédric Berney, Alastair Simpson, Philippe Deschamps
Proposed deadline = end of September 2016

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

Pending questions to address:

- Minimum amount of contextual data that we want to include/encourage
 - Synthetise a list based on ^{18S}EukRef guidelines and MlxS contextual data fields
 - Define which contextual data should apply to taxa vs. sequences to simplify integration of multiple gene markers into a single taxonomic framework
 - Present them in a guideline paper for protistologists
- Propose solutions to ensure compatibility between the UniEuk framework (unlimited number of ranks) and end-users needing a fixed number of ranks to operate (e.g. Vamps)

Implementation Group 4 - Taxonomy Guidelines / email: taxo-guidelines@unieuk.org

Suggested Participants:

Alastair Simpson, Sina Adl, Cédric Berney, Fabien Burki(?), Matthew Brown, Ramon Massana, Chris Lane(?), Laure Guillou, Sandra Baldauf, Laura Katz, Colomban de Vargas

Specific tasks already identified:

- Agree on the initial deep-branching structure of the taxonomy using phylogenomic evidence
Alastair Simpson, Sina Adl, Cédric Berney, etc.
Proposed deadline = end of September 2016

Pending questions to address:

- Work on the idea of a "UniEuk Taxonomic Council", liaising with and building upon such existing "councils" within international societies (ISOP / IPS)
- What happens when both Bot and Zool codes have been applied to the same clade?
- Propose appropriate guidelines whenever general taxonomic questions arise during the course of the project as more clades are "^{18S}EukRefed" and work on the taxonomy map proceeds

Implementation Group 5 - UniEuk Outreach / email: outreach@unieuk.org

Suggested Participants:

Pelin Yilmaz, Colomban de Vargas, Cédric Berney, Bremen implementer, Ginny Edgcomb, Eunsoo Kim, Line Le Gall, Javier del Campo, Stefan Geisen, David Caron, Jon Kaye and Sara Bender at the Moore Foundation.

Specific tasks already identified:

- Work on conceptual & practical case studies that can become a communication device to clarify the UniEuk mission and explain the workflow
 - Define personas of different stakeholders corresponding to specific use case scenarios
 - Write 1-2 pages narratives to explain how UniEuk will address them**Cédric Berney and various committee members**
proposed deadline = end of July 2016
- Website and other material for outreach and to present the project at conferences
 - Updated infographic, slides
 - Prepare a model of the website structure similar to that of eukref.org**Implementers and SC members**
Proposed deadline = end of September 2016

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Consolidate the list of stakeholders to (re-)contact (other initiatives, tools & databases that can benefit from UniEuk outputs, culture collections & museums, journals, potential funders, etc.)
 - Expand the initial list available in the google drive
 - Prior to (re-)contacting them, prepare type letters and determine clearly what we want to get from them and what benefit(s) they will get from liaising with UniEuk

SC and other committee members

proposed deadline = end of November 2016

- White paper presenting the project
 - Short text (2 pages maximum) explaining the rationale and goals of the project
 - One main figure summarizing the project's organisation & workflow (cf. updated infographic)

Colomban, Pelin, and Cedric will write a draft

proposed deadline = end of November 2016

Pending question to address:

- Create a UniEuk wish-list of resources, actions, side projects, etc. that we want to encourage in parallel to core UniEuk deliverables

Implementation Group 6 - Image Data / email: image-data@unieuk.org

Suggested Participants: **Sébastien Colin**, Colomban de Vargas, Pelin Yilmaz, Javier del Campo, Cédric Berney, Bremen implementer, Marc Picheral, STAB members

Pending questions to address:

- Determine the minimum image quality and contextual data associated with images that we want to encourage (e.g. magnification, scale bar or size, objective type, microscopy settings, etc.)
- Encourage contributors to deposit images in EcoTaxa and encourage Euro-BioImaging to consider hosting Ecotaxa in the long-term
- Identify and contact key people who may have useful personal collections of image data

Additional Tasks

1. Organization of the next ^{18S}EukRef workshops

Suggested Participants:

Javier del Campo, Laura Wegener-Parfrey, Colomban de Vargas, Matthew Brown, Cédric Berney, Ramon Massana, Raffaele Siano, Alastair Simpson, Sina Adl, Line Le Gall

email: eukref-workshops@unieuk.org

Specific tasks already identified:

- Next ^{18S}EukRef workshop in Barcelona week of Nov 28 to Dec 2, 2016
 - Confirm dates and start organising the workshop
 - Decide what main protist groups need to be covered and contact relevant advisers (e.g. David Bass, Matthew Brown, Enrique Lara, etc. in the case of Amoebozoa / Rhizaria)

Javier del Campo, Laura Wegener-Parfrey, Colomban de Vargas, Ramon Massana, etc.

proposed deadline = end of July 2016

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Possibility to hold a ^{18S}EukRef workshop at a dinoflagellate meeting in Bordeaux in July 2017
 - If so, a workshop proposal has to be submitted to the meeting organizers
 - Decide the scope of the workshop (dinoflagellates sensu lato / all non-ciliate alveolates?) and contact relevant advisers (e.g. Patrick Keeling, Laure Guillou, etc.)

Javier del Campo, Laura Wegener-Parfrey, Colombar de Vargas, Raffaele Siano

proposed deadline = end of July 2016

- Liaise with the Animal, Plant and Fungal communities to organise curation of 18S sequences in a third large ^{18S}EukRef workshop (autumn 2017?) - cf. list of names in Meeting Notes

Javier del Campo, Laura Wegener-Parfrey, Colombar de Vargas, Line Le Gall, etc.

proposed deadline = end of November 2016

2. Consolidating the Groups of Taxonomy Experts

Cédric Berney, SC and AC members

proposed deadline = end of November 2016

- Divide the protist tree of life into meaningful, workable units, propose 2-3 highly consensual lead experts for each group (C.B. and the SC), get feedback and approval from the AC
- (Re-)contact lead experts with a very clear description of what is expected from them (i.e. contact points of the Taxonomy Coordinator sharing editing rights on the taxonomy map + coordinators of the decision process in their group, but without higher decision power)
- Consolidate the groups of experts
 - Expand the initial lists of experts from the project proposal and contact all
 - Make sure all ^{18S}EukRef contributors are included in the relevant group(s) of experts

3. Identify and contact key additional people to invite to join the AC

SC and AC members

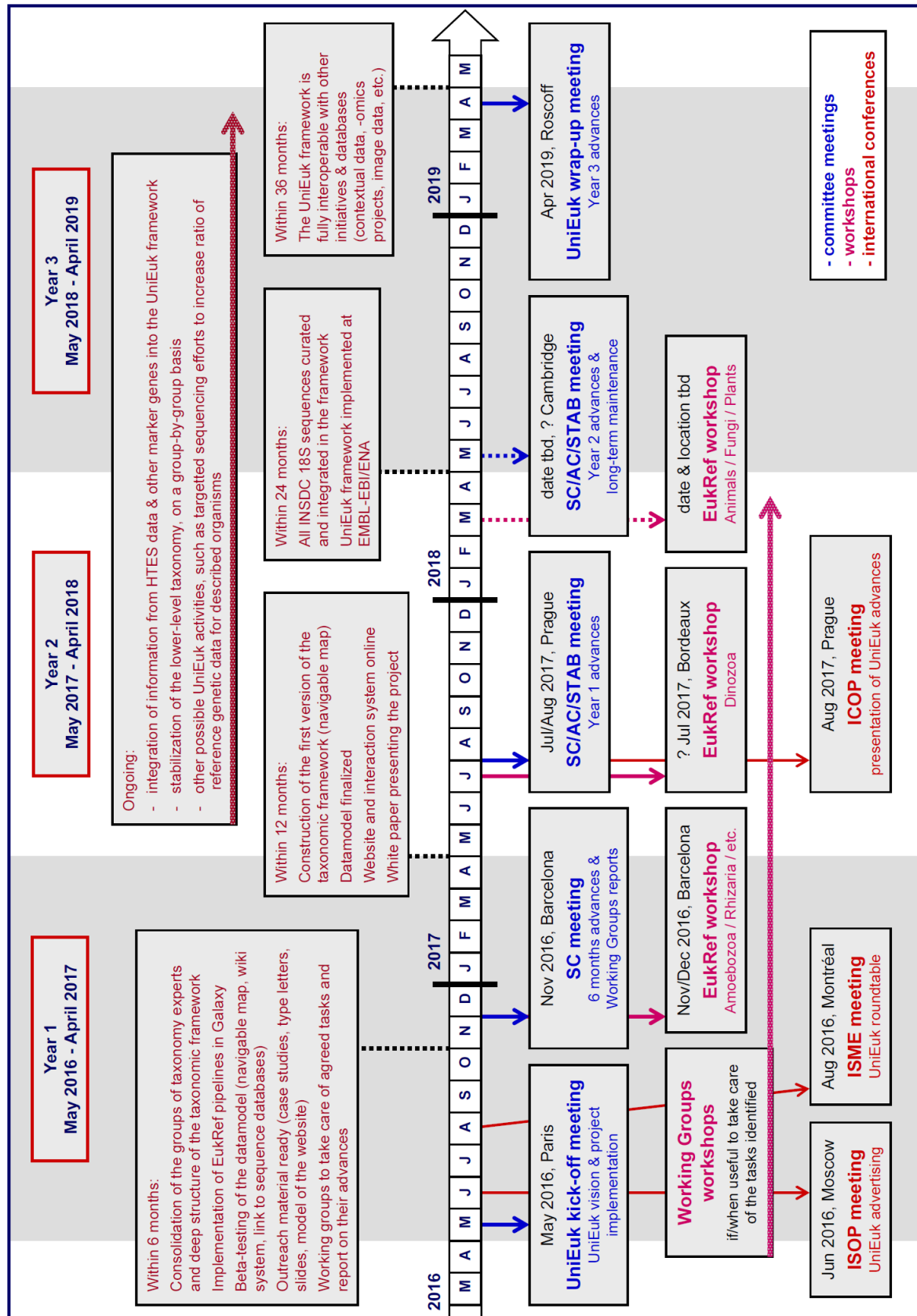
proposed deadline = end of November 2016

- Another phycologist / a mycologist (e.g. Tom Richards?) / a person involved in the NCBI taxonomy (e.g. Detlef Leipe?) / a person involved in the APG initiative / a person involved in global metazoan taxonomy (e.g. Gonzalo Giribet?)

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

D. UniEuk Timeline



E. Meeting Notes

Topic 1 - ^{18S}EukRef : where are we and what's next ?

(based on notes from Stéphane Audic, Javier del Campo, Sara Bender, ...)

Reminder:

^{18S}EukRef provides a semi-automated pipeline to capture sequences from cultured and uncultured eukaryotes belonging to a consistent and monophyletic group from INSDC databases and build reference alignments and trees. It was built with 18S sequences in mind. The pipeline starts with a user-provided set of reference 18S sequences, used for similarity search against the INSDC databases using Blastn. Newly retrieved sequences are examined to detect and remove chimeras, short sequences (<500bp), and sequences obviously outside the group of interest. Valid sequences are blasted again until no more sequences can be added. Then an alignment and a tree are built, using appropriate ingroup and outgroup sequences. Finally, taxonomic annotation is performed in a spreadsheet; environmental or contextual data can be retrieved from the original INSDC entries and used to inform the annotation process. Everyone is welcome to review, comment and improve the scripts. All details can be found on the eukref.org website.

Updates:

- The first ^{18S}EukRef workshop (Keeling's lab, UBC, Vancouver, July 2015) included many hands on sessions and was a success; part of the work done there is on the way to be published
- Curation of Ciliates and Excavates is almost finished
- Information available on the ^{18S}EukRef webpages: all scripts; guidelines on how to use them; list of all taxonomic groups (based on Adl et al. 2012) and who chose to curate them; guidelines on how to annotate sequences (especially environmental clades)
- Detection of sequences outside the group of interest is currently done by similarity search against Silva; as UniEuk progresses and its taxonomy is implemented in Silva, that part of the procedure is expected to become more consistent and efficient
- Convergence of the pipeline is observed in most cases after a few iterations; rarely the procedure "explodes" to include sequences from outside groups - if this happen, it is suggested to be mostly because of undetected chimeras

Decisions Taken:

- The ^{18S}EukRef pipeline will be implemented in the Galaxy instance already configured at the Station Biologique de Roscoff; this will offer the service to a community of users not familiar with software to run command lines / without the necessary computing resources
- ^{18S}EukRef outputs (annotated phylogenetic trees) will form the backbone of UniEuk
- Individuals actively working with the UniEuk taxonomic map/wiki for a given group should include all those who participated in the ^{18S}EukRef curation effort of that part of the tree
- A second general ^{18S}EukRef workshop could take place in Barcelona week of Nov 28 to Dec 2, 2016; two main groups to focus on could be Amoebozoa and Rhizaria
- For the purpose of completeness of the Ref 18S database to annotate HTES data, animals, plants and fungi have to be curated even if 18S is not the best marker for these groups

Pending Questions:

- Concerns were raised about the use of uchime; replacement by vsearch was recommended

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Where does ^{18S}EukRef stop and UniEuk start?
 - Think of ^{18S}EukRef as “T-zero” of UniEuk, where changes can be tracked
- How can ^{18S}EukRef outputs be used as input for UniEuk; how inclusive do they have to be?
- ^{18S}EukRef being clade-based, curation of sequences outside known clades must be addressed
- We need to be wary of seeking full consensus on the tree
 - When do we decide to proceed without full agreement?
- Take pending decisions about next workshops and how many more are needed
 - Determine what protist groups still need work done
 - A dinoflagellate-specific workshop could be organised during the 11th International Conference on modern and fossil Dinoflagellates, Bordeaux, July 2017
- Timing of integration of ^{18S}EukRef outputs / UniEuk taxonomy into Silva
 - Users worldwide rely on the database without questioning it, so annotation of environmental data and comparability will significantly improve as soon as Silva starts being updated
 - Compromise between having it right/exhaustively and having it quickly

Next Steps:

1. Check the ^{18S}EukRef pipelines and send comments by 31 May 2016
Frédéric Mahé, Chris Quast, Lucas Czech, Oleksii Kozlov
2. Implement ^{18S}EukRef pipelines in the Galaxy instance at SBR before end of summer 2016
? Stéphane Audic
3. Make sure all ^{18S}EukRef contributors are invited to participate in UniEuk
Cédric Berney
4. If we try to hold a dinoflagellate-specific ^{18S}EukRef workshop in Bordeaux in July 2017, a workshop proposal has to be submitted to the meeting organizers
? Raffaele Siano
5. Liaise with the Animal, Plant and Fungal communities to organise curation of 18S sequences in future ^{18S}EukRef workshops (e.g. Gonzalo Giribet, Jordi Paps, Iñaki Ruiz-Trillo for large scale metazoan phylogeny, Hans Helder for nematodes, Tom Richards for basal fungi, etc.)
? who ?

Topic 2 - UniEuk outputs & resources: community needs and expectations

(based on notes from Cédric Berney, Lucas Czech, ...)

Reminders:

- It is important to make the distinction between
 - **phylogenetic trees** related to the ^{18S}EukRef pipelines & EPA tools = resources shared with the community (with accompanying sequence alignments) + source of information for the UniEuk taxonomy (initially mostly 18S trees but later could be based on other markers)
 - **taxonomy map** = UniEuk taxonomy summary tree, not phylogenetic but schematic, composite of phylogeny- and morphology-based information
- Navigable tree = visualization tool of the consensus taxonomic framework (taxonomy map)
 - will include described, unsequenced taxa and sequenced, undescribed environmental OTUs
 - links between purely morphological and purely genetic lineages will be created whenever that information becomes available

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Possible practical output of ^{18S}EukRef/UniEuk = standardized phylogenetic placement of HTES sequences for taxonomic annotation and to further inform the UniEuk taxonomy

A. Swarm Bank

Updates & Decisions Taken:

- Low computational resources / storing capacity needed, so no technical restrictions
- V4 = priority because phylogenetically informative (cf. main goal = inform UniEuk taxonomy)

Pending Questions & Next Steps:

- Empirical tests of merging different V4 datasets into a single set of swarms

Frédéric Mahé

- Online public interface, exploration and visualization of the Swarm Bank
 - What do we want users to be able to do?
 - How to provide the information associated with the raw data? (origin etc.)
 - Useful output would be to allow comparisons between primers / extraction methods used

B. Evolutionary Placement Algorithm and relevant applications

Updates & Decisions Taken:

- Computational limitation: for now, 100,000 to 200,000 reference sequences (alignment & corresponding tree) are a realistic limit; placement of short reads then follows a linear increase so there's no computational limitation there
- A multi-level placement ("Russian doll") approach will be useful to perform taxonomic annotation with EPA, with reference alignments and trees at different taxonomic levels
 1. EPA run against a coarse, eukaryote-wide reference tree to sort HTES sequences into clades
 - SILVA = good starting point because eukaryote-wide alignment & tree already available
 2. EPA run against finer, more detailed, group-specific trees to get a precise annotation
 - Use available ^{18S}EukRef group-specific alignments & trees

Pending Questions & Next Steps:

- Is placement of HTES reads into phylogenetic trees to inform the taxonomy framework really part of UniEuk or just a side project? cf. time and computational limitations

C. Interaction with the community: UniEuk navigable tree & wiki system

Updates & Decisions Taken:

- The UniEuk taxonomy map must be seen as work-in-progress at all times
 - Will continually improve over time / will always remain perfectible
 - Using information from environmental data is a particularity of our system but implies facing a huge unnamed diversity and placing it in the taxonomic framework
 - Need for regular frozen versions for dissemination to INSDC and gene reference databases
- Suggestion = 3 levels in the interaction loop
 - Anybody can register to navigate and comment on the taxonomy map
 - Groups of experts are created for meaningful taxonomic units (as many as necessary to cover the whole diversity of protists) and will be in charge of validating modifications
 - 2/3 "lead experts" in each group are the main contact points of the Taxonomy Coordinator and share direct editing rights with him
- Useful to have a way to flag nodes in the taxonomy map according to what each taxon entails

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Morphol. data available vs. environmental clade / molec. data available vs. unsequenced organism / phylogenetically supported clade vs. evolutionary hypothesis / etc.
- Possibility to navigate corresponding subsets of taxa/nodes if technically feasible
- Navigable tree: computational limitations suggest that a "Russian doll" approach is necessary
 - Open Tree of Life model : limited number of "ranks" visible at any time / loads subtrees as you navigate the tree / shows full path to root at any point / advantage = comment section
 - Ability to correct mistakes easily and relatively quickly is important
 - Other approaches to keep in mind : graph database / fractal visualization

Pending Questions & Next Steps:

- Divide the tree of life into meaningful units / consolidate the groups of experts
Cédric Berney and the SC

D. UniEuk / ^{18S}EukRef website

Updates & Decisions Taken:

- Online presence and contact email address are needed soon now that the project has started
 - Domain unieuk.org has been reserved
- All content must be agreed upon internally (e.g. via the google drive) before going online
 - Possible website content: project proposal with revised infographics of the project structure / committee members & roles of the committees / UniEuk talking points, case studies & other material to present the project / meeting reports / UniEuk calendar / etc.
- Implementation deadline: follow the motto "under-promise, over-deliver"
 - We don't want to advertise what we will offer until relevant practical questions have been answered and implementation of the work modules has been agreed upon

Pending Questions & Next Steps:

- Prepare a model of the website structure similar to that of eukref.org
- Work on conceptual & practical case studies and other material to present the project online
Cédric Berney and various committee members

Topic 3 - The UniEuk datamodel and contextual data

(based on notes from Pelin Yilmaz, Christian Quast, ...)

This session included three major discussion points, which based on prior discussions were deemed important to determine the data model for the UniEuk system.

A. Contextual data : deciding what we need & agreeing on a common vocabulary

During discussion of this point, we first clarified whether it was UniEuk's responsibility to collect relevant contextual data. We quickly agreed that this would be out of UniEuk's scope, as UniEuk will be a "secondary database", and should instead harvest available and relevant contextual data from primary data repositories, such as INSDC. We continued this point with a discussion of what contextual metadata is relevant for UniEuk. Pelin Yilmaz, Javier del Campo, and Colombar de Vargas presented options from existing standards and projects, and the participants suggested additional fields to be included in such a minimal UniEuk contextual data collection. It was quickly obvious that this topic required some more extensive and offline discussion, therefore, for the sake of time-keeping, we agreed that a proposal of minimal contextual data fields will be synthesized out of existing ^{18S}EukRef fields, plus additional MIxS contextual data fields. Another

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

important outcome of this discussion was the agreement to write a short guideline paper for protistologists, which explains, from a practical perspective, how to accurately fill in such contextual data sheets. Finally, we also discussed the necessity of applying separate contextual data to taxa and to sequences. We concluded that this would simplify the integration of multiple gene markers into a single taxonomic framework.

B. Not reinventing the wheel : co-opting existing systems for multicellular eukaryotes

Discussion about the taxonomic frameworks that would be adopted by UniEuk for plants, animals, and fungi. We discussed the best initiatives to contact for different parts of the tree that would not be initially covered by UniEuk (cf. parts of Discussion Topic 5). These were:

Fungi: Unite / Tom Richards (Exeter)

Algae: Algaebase

Higher plants: NCBI taxonomy / Angiosperm Phylogeny Group (APG) initiative

Animals: NCBI taxonomy / Open Tree of Life

The discussion strayed from the main direction and we also discussed whether curating the taxonomy of the above-mentioned groups should also be part of the UniEuk aims. This was ultimately deemed out of scope, but we did note, in the future, we should encourage scientists interested in these groups to join the UniEuk portal, and help curating these parts of the eukaryotic tree of life. For the sake of not antagonizing any particular group, we have agreed to adopt the NCBI taxonomy by default for plants, animals, and fungi, at least for a short while.

C. Beyond 18S: integration of other reference databases in the system

Discussion of this point was brief. The group listed the most-wanted alternative marker genes (28S rRNA, ITS, COX, plastidial 16S rRNA, fragments of ribosomal genes - V4, V9, end of 18S to beginning of 28S across ITS), and Christian Quast pointed out that from a database technical perspective, as many alternative markers can be incorporated as desired. An important point arose from this - namely linking entries of different genes to the same taxon. This of course requires confirmation that two (or more) gene sequences originate from the exact same isolate. A few solutions to this problem were suggested:

- transcriptomics / genomics data > contig-style accession numbers (AZXO000001) make it straightforward to link different markers
- 18S / ITS / 28S from same culture collection strain / isolate name
- list correspondence between taxID and accession numbers
- straininfo tracks info from culture collections worldwide, useful resource to connect with

Also in relation to point C, we discussed further integration issues, for example integrating morphology and genetic data. We agreed that image data hosting had to be outsourced (for example to EcoTaxa - <http://ecotaxa.sb-roscoff.fr/>). We also agreed to liaise with journals, and recommend them that they encourage images (or links to them) being placed in public repositories, together with sequences. We concluded this topic with a discussion on how to best represent links to morphological data in the database model, and following Christian Quast's suggestion, agreed that the usual method of DOI/URL linking will be sufficient.

Other notes in relation to that discussion:

We spent the rest of the allocated discussion time on more technical topics. Christian Quast suggested creating a mock-up UniEuk interface, to be appraised by 3-4 core testers. This suggestion was well-received, and we agreed that Alastair Simpson, Cédric Berney, and Philippe Deschamps should be the testers. The mock-up interface will be created by Christian Quast, and he will take UNITE and Open Tree of Life interfaces into account.

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

We concluded our session with clarifications on a few technical points, specifically:

- Data storage: relying on Christian Quast's expertise
- RESTFUL webservice, API interface: will be provided for power users
- Public code hosting: will be done via bitbucket or github
- Jira can constitute the ticketing system for the taxonomy discussion system
- The possibility to store ^{18S}EukRef outputs in SILVA, which will depend on the ability to update and maintain the data / this is an open discussion topic

Summary of Decisions Taken & Next Steps:

- Providing an interface to fill contextual data is outside the scope of UniEuk; all metadata should be automatically imported from INSDC / ^{18S}EukRef outputs
- Usefulness of separating contextual data applying to taxa vs. sequences to simplify integration of multiple gene markers into a single taxonomic framework
- Synthesise a proposal of minimal contextual data fields out of existing ^{18S}EukRef and MIxS contextual data fields
- Write a contextual data guidelines paper for protistologists
- By default, initially adopt the NCBI taxonomy for plants, animals, and fungi
- Most useful alternative genetic markers: 28S rRNA, ITS, COX, plastidial 16S rRNA, fragments of ribosomal genes (V4, V9, end of 18S to beginning of 28S across ITS)
- Key point = linking entries of different gene sequences to the same taxon when known
- Mock-up interface to be created and appraised by 3-4 testers

Christian Quast / testers: Cédric Berney, Alastair Simpson, Philippe Deschamps

Topic 4 - Protist taxonomy guidelines in the -omics era

(based on notes from Cédric Berney, Sina Adl, ...)

Reminder:

UniEuk aims to build a universal taxonomic framework for eukaryotes (specifically protists); the resulting taxonomy will be robust and integrative, yet flexible and consensual, with clear practical goals in mind: it will inform databases of validated genetic markers and be applied to sequence data held within the INSDC repositories, for the benefit of the broader community.

- The project is not a vehicle where people can hash out very specific issues of taxonomy
- Discussing the species concept in protists, discussing the pros & cons of the Bot. and Zool. codes, and/or promoting a "protist nomenclature code" are all beyond the scope of UniEuk

Updates & Decisions Taken:

- Starting point = existing taxonomy guidelines drafted during the first ^{18S}EukRef workshop, especially in the case of environmental clades
- A variable number of levels in the taxonomy is necessary, but classical ranks can be retained where appropriate for informative reasons
- The default approach to name taxa at supra-species levels should be very conservative
 - Only include phylogenetically supported nodes useful to the purpose of the UniEuk taxonomy
 - Polytomies are accepted/encouraged whenever it's not useful to consider finer separations
- Specific case of polyphyletic genera (e.g. *Gymnodinium*): distinguish lineages using guidelines similar to those for environmental clades (e.g. *Gymnodinium_1*, *Gymnodinium_2*, etc.)
- "Use your best judgement" as general guideline to all groups of experts

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Integrative approach: recommending multiple congruent lines of evidence to justify clades
 - When genetic markers disagree: follow the marker considered most appropriate at this level
- Agreement that a "UniEuk Taxonomic Council" would be useful as a long-term legacy of UniEuk beyond 3 years and as a buffer above group-specific councils to avoid cliques
 - Build upon such existing "councils" within ISOP / IPS
 - Needs to include odd number of highly consensus-driven people!
- It is not in our remit to formally rename clades / create taxa, but UniEuk will represent a way to flag problematic situations to the relevant community

Pending Questions & Next Steps:

- Suggestion to go down to isolate/strain as lowest taxonomic unit (whenever possible)
 - To facilitate identification of issues (wrong species identifications, cryptic species, etc.)
 - To ensure links between genetic markers only happen when from the exact same isolate
- What happens when both the Bot. and Zool. codes have been applied to the same clade?
- Release frequency (versioning of the navigable taxonomy map): 6 months?
- Once more clades have been "EukRefed" and we have a reasonably good starting version of the taxonomy, potential issues warranting more group decisions will be easier to identify

Topic 5 - Interoperability of UniEuk with other initiatives

(based on notes from Pelin Yilmaz, Cédric Berney, ...)

This session focused on the following discussion points:

- Which existing initiatives do we want to liaise with (taxonomy, genetic & image databases)?
- Are there any specific requirements that can be implemented from the start to ensure that UniEuk outputs are compatible with the expectations / needs of these initiatives?

Reminder: UniEuk is not a database of images or contextual data, therefore we want to create links between taxa and relevant entries in external initiatives / databases

Updates & Decisions Taken:

- Prior to contacting other initiatives / databases, determine and be ready to explain clearly
 - **what we want to get from them**
(e.g. taxonomic information, contextual data, link to images, scientific endorsement)
 - **what benefits they will get from liaising with UniEuk**
(e.g. added value to their content from using a unified taxonomy / cross-linking with other initiatives, access to an exportable, interactive UniEuk taxonomy map)
- Follow the general guiding principle of "**under-promise, over-deliver**" / have appropriate communication material ready before contacting other initiatives (talking points, website, ...)
- Image data: our priority is well-curated images of sequenced and/or type species
 - Taxonomic resolution: images have to enhance the taxonomy to be useful
- Having a system with **unique taxonomic IDs** for each node in the taxonomic framework will in almost all cases be sufficient to communicate with other initiatives
(e.g. links to INSDC accession numbers, publications, image databases, etc.)
- We need to strongly encourage systematic links between image and sequence data (especially in newly described organisms) / liaise with journals so that they encourage it too
- Links to citizen science initiatives is a consideration, as a means to involve museums and increase buy-in to UniEuk

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

Initiatives/Databases already linked / interested to link with UniEuk:

- EMBL/EBI-ENA → project partner ensuring link to INSDC + access to Euro-BiolImaging
- SILVA / PR2 → curated 18S reference databases benefitting from the ^{18S}EukRef outputs (reference 18S alignments and trees) and the UniEuk taxonomic framework
- EcoTaxa → image database already committed to using the UniEuk taxonomy
 - Capability to include taxa that don't have sequences yet / existence of image data can be used to determine priority for inclusion in the taxonomic framework
 - Encourage contributors to deposit images in EcoTaxa
 - Encourage Euro-BiolImaging to consider hosting Ecotaxa in the long-term
- VAMPS → visualization and analysis of microbial population structures
 - As much contextual data as possible important to facilitate meta-analyses on env. datasets
 - Fixed number of taxonomic ranks is a necessity
- Pangaea → visualization and analysis of microbial population structures
 - Only need unique UniEuk taxonomic IDs they can point to

Other Initiatives/Databases to contact/liase with in priority and why:

- General databases: Open Tree of Life, Catalogue of Life, Encyclopedia of Life, Worms, etc.
 - Source of general taxonomic information
 - Can benefit from our curated protist taxonomy / cross-links with other databases
- Algaebase (currently maintained with in-kind contributions)
 - Source of taxonomic information for algae
 - Determine if UniEuk can help provide longer-term survival/hosting of Algaebase?
- Fungal initiatives: Unite, FUNGuild, VAMPS/MoBeDAC, Jason Stajich/FungiDB, Paul Kirk
 - We can benefit/learn from the taxonomic rules used in the fungal community
- Group-specific: Ciliate Research Coordination Network, ForamBarcoding, VIMS, etc.
 - Potential source of specific taxonomic information, images and contextual data
- Culture collections (CCAP, ATCC, RCC, NCMA-Bigelow), Museums, MMETSP & single-cell genome/transcriptome projects, Micro*scope (precursor to EoL), etc.
 - Institutions/projects likely to have rich environmental data associated with images/taxa
- Others? (CBOL/BOLD, GNITE, UNION, NSF Genealogy of Life, EukPathDB, etc.)

Pending Questions & Next Steps:

- Define the minimum image quality and associated contextual data that we want to encourage
 - e.g. magnification, scale bar or size, source location, objective type, microscopy settings, preservation method, type of imaging platform (+ depth, altitude, elevation?)
- The UniEuk taxonomic framework must be designed to accommodate an unlimited number of ranks, but some end-users / project partners (e.g. PR2, Vamps) currently have and/or need a fixed number of ranks to operate
- Prepare different type letters to contact different stakeholders / taxonomy initiatives / image databases / culture collections / museums & botanical gardens / animal, fungal & plant communities / individual contributors / etc.
- Identify key people to contact (personal collection of image data, leaders of relevant initiatives) - e.g. Ferry Siemensma, Eckhard Voelker, Paddy Patterson, John Dolan, O. Roger Anderson, Alexey Smirnov, Weibo Song, Mike Paige, Dennis Lynn, etc.

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

- Identify all available tools that can benefit from the UniEuk outputs
- e.g. Qiime, Qiitta, Mothur, Megan, MG-Rast, etc.
- Try to have representation in meetings outside of protistology/phycozoology, such as ASLO, Ocean Sciences, Ecological Society of America, ICSZ
- Create a UniEuk wish-list of resources, actions, side projects, etc. that we want to encourage in parallel to core UniEuk deliverables

Topic 6 - UniEuk beyond Year 1 - outreach, fundraising, etc.

(based on notes from Cédric Berney, Javier del Campo, ...)

Community outreach / UniEuk talking points when engaging community members

Updates & Decisions Taken:

- An initial list of various potential stakeholders was built by Cédric Berney and is available in the google drive; it indicates who has already been contacted / has responded
- UniEuk has already been advertised in the following meetings / symposia / workshops:
 - Jul 2015 1st ^{18S}EukRef workshop, Vancouver, Canada
 - Sep 2015 ISOP / ECOP joint meeting, Sevilla, Spain
 - Oct 2015 Oceanomics / TARA annual meeting, Roscoff, France
 - Jan 2016 EMBO/EMBL Symposium: A New Age of Discovery for Aquatic Microeukaryotes, Heidelberg, Germany
 - Feb 2016 Deutsche Gesellschaft für Protozoologie annual meeting, Saignelégier, CH
 - Apr 2016 Protistology UK (ex-BSPB) annual meeting, Bournemouth, UK
- The next wave of outreach to potential stakeholders must be **goal-oriented**; we need to explain clearly what we expect from them and how they will benefit from UniEuk
- When communicating, we need to be proactive and adapt the communication strategy based on the nature of the relationship we want to establish with a particular stakeholder
- We will prepare use case scenarios that will become a communication device to clarify the UniEuk mission, help explain how UniEuk is integrative, what it is, what it is not
 - Everyone internally will have examples of how the UniEuk pipeline works
 - External stakeholders will understand the UniEuk mission, structure, and implementation
- The Angiosperm Phylogeny Group (APG) initiative may provide insight into structure and language (for discussing UniEuk) as it had a similar mission

Pending Questions & Next Steps:

- More potential stakeholders should be added to the existing list (e.g. Catalogue of Life, Gbif)
 - Main categories of stakeholders to consider: 1. end-users facilities & databases / 2. potential funders / 3. journals, societies, museums, culture collections, etc.

All committee members can suggest additional stakeholders to contact and why
- Case studies: define a few use case scenarios and develop them into 1-2 page narratives
Cédric Berney and various committee members
- Prepare and share appropriate communication material approved by all
 - **Set of UniEuk slides** that we can all use to present / advertise the project at meetings, etc.
 - Different type letters to contact different kinds of stakeholders
SC and other committee members

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

White paper presenting the project (PLOS Community Page?)

- Importance of having a citable reference for the project relatively early on was discussed
- One main figure summarizing the project's organisation and workflow
 - Project's inputs (^{18S}EukRef outputs / genetic & morphol. knowledge / community expertise)
 - Project's output (the universal taxonomic framework implemented in EMBL/EBI-ENA)
 - Main links to initiatives / databases / stakeholders / end-users benefitting from UniEuk
- Short text (2 pages maximum) explaining the rationale and goals of the project to the broad scientific community (cf. UniEuk vision & what UniEuk is / is not)

Colomban de Vargas and Pelin Yilmaz will write a draft when the time is right

Fundraising for years 2-3 and UniEuk-related activities

- Determine potential funding sources and share information about funding opportunities
 - Create a document in the google drive listing all possibilities

The SC oversees fundraising but concrete suggestions welcome from all

- Potential funding sources mentioned:
 - Simons Foundation - SCOPE - Funding raised by the taxonomy experts
 - NSF International Collaboration Educational Networks - NSF RCN
 - Agnes b. (Tara Oceans) - Funding from regional funding agencies

Project timeline and milestones, data release checkpoints, scheduling the next meetings

Updates & Decisions Taken:

- Pelin Yilmaz successfully submitted a proposal for a UniEuk roundtable at the 16th ISME meeting in Montréal, Canada, Aug 2016 - <http://www.isme-microbes.org/isme16>
- General agreement: ICOP 15 meeting in Prague, Czech Republic, Jul/Aug 2017 = ideal venue for a main UniEuk committees meeting / to advertise the first results - <http://www.icop2017.org/>
- Next ^{18S}EukRef workshop proposed to be held in Barcelona in Nov 2016
- Possibility to submit a proposal for a ^{18S}EukRef curation / UniEuk information workshop on dinoflagellates at the 11th International Conference on modern and fossil Dinoflagellates, Bordeaux, France, Jul 2017
- Instead of more big implementation meetings, proposal to organise several smaller meetings (up to 10 people max.?) centred on specific work packages agreed upon by the committees

Pending Questions & Next Steps:

- Different UniEuk/^{18S}EukRef outputs (curated 18S database / consensus taxonomic framework) may interest different end-users, with a potentially different release policy

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

F. Meeting Agenda

- 1. Meeting Goals:** Take practical decisions about the implementation of the UniEuk project, how and when key tools and stages of the project should take place, refine the planned outcomes and community resources, so that implementers have a clear list of goals and deliverables to achieve, especially during the first 6 to 12 months of the project.

- 2. Summary of the Meeting Programme:**

Day 1

- Welcome address by the Directors of the project, Colombar de Vargas & Pelin Yilmaz
- Overview of the roles and responsibilities of the UniEuk committees
 - **Steering Committee**
deciding body of the project, responsible for the communication strategy, additional fundraising, and overseeing the project's progress; composed of members with complementary expertise covering all aspects of the project (from protist taxonomy to genetic reference databases).
 - **Scientific & Technical Advisory Board**
direct source of help and advice in relation to the project's implementation; composed of experts in genetic reference database management, eukaryote taxonomy & barcoding, and analysis of large high-throughput environmental sequencing datasets.
 - **Advisory Council**
ensures the scientific leadership, credibility and promotion of the project, as well as community buy-in and representation of the main stakeholders; composed of key figures in protistology.
- Introduction of all the meeting's participants and their role in UniEuk
- Outline of the programme and organisation of the meeting
- General introduction and presentation of the discussion topics via short talks
 - *Reference gene databases: state-of-the-art and challenges* - Pelin Yilmaz
 - *Eukaryote taxonomy: state-of-the-art and challenges* - Alastair Simpson
 - *The Gordon & Betty Moore Foundation and how UniEuk fits into their science program*
Sara Bender
 - *The ^{18S}EukRef initiative and 1st ^{18S}EukRef workshop* - Javier del Campo
 - *Summary of the proposed UniEuk workflow and planned outputs* - Cédric Berney
 - *The Swarm clustering approach and UniEuk Swarm Bank* - Frédéric Mahé
 - *The Evolutionary Placement Algorithm (EPA) and possible applications* - Lucas Czech
 - *EcoTaxa, online visual exploration & taxonomic annotation of eukaryote images*
Marc Picheral
- Parallel Discussion Session I – **Topic 1 - ^{18S}EukRef : where are we and what's next ?**
chair: Javier del Campo reporter: Stéphane Audic
attending: Colombar de Vargas, Eunsoo Kim, Frank Oliver Glöckner, Alastair Simpson, Micah Dunthorn, Frédéric Mahé, David Caron, Sara Bender, Christian Quast, Albert Reñé, Lucas Czech, Oleksii Kozlov, Philippe Deschamps, Guifré Torruella
- Parallel Discussion Session I – **Topic 5 - Interoperability of UniEuk with other initiatives**
chair: Pelin Yilmaz reporter: David Bass

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

attending: Ginny Edgcomb, Sina Adl, Stefan Geisen, Linda Amaral Zettler, Line Le Gall, Connie Lovejoy, Cédric Berney, Jon Kaye, Jeena Rajan, Maria Holzmann, Stéphane Pesant, Marc Picheral

Day 2

- Reporting-back session on the first Parallel Discussion Session by the reporters (Stéphane Audic & David Bass)
- General introduction and presentation of the discussion topics via short talks (continued)
 - *EMBL/EBI-ENA and its involvement in UniEuk* - Guy Cochrane
 - *The UniEuk datamodel - simplicity, flexibility, efficiency* - Christian Quast
 - *VAMPS, visualization and analysis of microbial population structures* - Linda Amaral Zettler
 - *The Soil Protists Initiative* - Stefan Geisen
- General Discussion Session I –
Topic 2 - UniEuk outputs & resources: community needs and expectations
chairs: Colomban de Vargas & Pelin Yilmaz reporter: Cédric Berney (all attending)
- Parallel Discussion Session II – **Topic 3 - The UniEuk datamodel and contextual data**
chair: Pelin Yilmaz reporter: Christian Quast
also attending: Colomban de Vargas, Ginny Edgcomb, Guy Cochrane, Javier del Campo, Stefan Geisen, Frank Oliver Glöckner, Stéphane Audic, Frédéric Mahé, Sara Bender, Jeena Rajan, Lucas Czech, Oleksii Kozlov, Stéphane Pesant, Philippe Deschamps
- Parallel Discussion Session II – **Topic 4 - Protist taxonomy guidelines in the -omics era**
chair: Alastair Simpson reporter: Sina Adl
also attending: Eunsoo Kim, Linda Amaral Zettler, Line Le Gall, Micah Dunthorn, Connie Lovejoy, David Bass, David Caron, Cédric Berney, Jon Kaye, Maria Holzmann, Albert Reñé, Guifré Torruella

Day 3

- Reporting-back session on the second Parallel Discussion Session by the reporters (Christian Quast & Sina Adl)
- **Agreeing on common UniEuk principles**

- <i>UniEuk Vision, Mission, & Guiding Principles</i>	Sara Bender
- <i>What is UniEuk? What is not UniEuk?</i>	Pelin Yilmaz
- <i>UniEuk inputs and outputs, flow chart</i>	Colomban de Vargas, Javier del Campo
- <i>Project case studies</i>	Cédric Berney
- <i>Decisions and topics for future discussion</i>	Javier del Campo
- <i>Roles and Responsibilities, Project Timeline</i>	Colomban de Vargas
- General Discussion Session II – **Topic 6 - UniEuk beyond Year 1 - outreach, fundraising, etc.**
chairs: Pelin Yilmaz & Colomban de Vargas reporter: Eunsoo Kim
(all attending)
- Closing remarks by the Directors of the project, Colomban de Vargas & Pelin Yilmaz

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

G. Meeting Participants List

Colomban de Vargas	UniEuk Director	CNRS - Station Biologique de Roscoff, France
Pelin Yilmaz	UniEuk Co-Director	MPI & Jacobs University, Bremen, Germany
Sina Adl	Steering Committee	University of Saskatchewan, Canada
Guy Cochrane	Steering Committee	EMBL-EBI / ENA, Cambridge, UK
Javier del Campo	Steering Committee	University of British Columbia, Canada
Virginia Edgcomb	Steering Committee	Woods Hole Oceanographic Institution, USA
Stefan Geisen	Steering Committee	Netherlands Institute of Ecology, Netherlands
Frank Oliver Glöckner	Steering Committee	MPI & Jacobs University, Bremen, Germany
Eunsoo Kim	Steering Committee	American Museum of Natural History, NYC, USA
Alastair Simpson	Steering Committee	Dalhousie University, Halifax, Canada
Linda Amaral Zettler	Scientific & Technical Advisory Board	Woods Hole Marine Biological Laboratory, USA
Line Le Gall	Scientific & Technical Advisory Board	Museum National d'Histoire Naturelle, Paris, France
Stéphane Audic	Scientific & Technical Advisory Board	CNRS - Station Biologique de Roscoff, France
Micah Dunthorn	Scientific & Technical Advisory Board	University of Kaiserslautern, Germany
Frédéric Mahé	Scientific & Technical Advisory Board	CIRAD, Montpellier, France
Connie Lovejoy	Advisory Council	Université Laval, Québec, Canada
David Bass	Advisory Council	Natural History Museum & CEFAS, UK
David Caron	Advisory Council	University of Southern California, USA
Cédric Berney	UniEuk Taxonomy Coordinator	UPMC - Station Biologique de Roscoff, France
Jon Kaye		Gordon & Betty Moore Foundation, USA
Sara Bender		Gordon & Betty Moore Foundation, USA
Jeena Rajan	Cochrane team	EMBL-EBI / ENA, Cambridge, UK
Christian Quast	Glöckner team	MPI & Jacobs University, Bremen, Germany
Maria Holzmann	Pawlowski team	University of Geneva, Switzerland
Albert Reñé	Massana team	Institut de Ciències del Mar, Barcelona, Spain
Lucas Czech	Stamatakis team	Heidelberg Institute for Theoretical Studies, Germany
Oleksii Kozlov	Stamatakis team	Heidelberg Institute for Theoretical Studies, Germany
Stéphane Pesant	PANGAEA database	Marum Centre for Marine Env. Sciences, Germany
Marc Picheral	EcoTaxa database	Laboratoire d'Océanographie de Villefranche, France
Purificación López-García		Université Paris-Sud, France
Philippe Deschamps	López-García team	Université Paris-Sud, France
Guifré Torruella i Cortés	López-García team	Université Paris-Sud, France

UniEuk 1st Implementation Meeting Report

UniEuk Kick-off meeting, 3-5 May 2016, MNHN, Jardin des Plantes, Paris

H. UniEuk Committees

Steering Committee (SC)

Colomban de Vargas

France, CNRS - Station Biologique de Roscoff, Roscoff – BioMarKs, Tara-Oceans, & CBOL ProWG, coordinator / *UniEuk* director

Pelin Yilmaz

Germany, Max Planck Institute & Jacobs University, Bremen – chair of SILVA's eukaryote taxonomy working group (ETWG) / *UniEuk* co-director

Sina Adl

Canada, University of Saskatchewan, Saskatoon – leader of ISOP's universal eukaryotic taxonomic framework

Juliet Brodie

UK, The Natural History Museum, London – board member of the International Phycological Society

Guy Cochrane

UK, EMBL-EBI / ENA, Cambridge – head of the European Nucleotide Archive at EMBL-EBI

Javier del Campo

Canada, University of British Columbia, Vancouver – EukRef coordinator

Virginia Edgcomb

USA, Woods Hole Oceanographic Institution, Woods Hole MA – board member of the International Society of Protistologists

Stefan Geisen

The Netherlands, Netherlands Institute of Ecology, Wageningen – leader of the Soil Protists Initiative

Frank Oliver Glöckner

Germany, Max Planck Institute & Jacobs University, Bremen – head of SILVA and MicroB3/OSD

Eunsoo Kim

USA, American Museum of Natural History, New York City NY – curator of Protistology at the AMNH

Alastair Simpson

Canada, Dalhousie University, Halifax – chair of ISOP's protist systematics committee

Advisory Council (AC) - TO BE COMPLETED

David Caron (AC chair)

Sandra Baldauf

Sonya Dyhrman

Laura Katz

Connie Lovejoy

Alexandra Worden

John Archibald

David Bass

Patrick Keeling

Jan Pawlowski

University of Southern California, Los Angeles CA, USA

University of Uppsala, Uppsala, Sweden

Columbia University, New York City NY, USA

Smith College, Northampton MA, USA

Université Laval, Québec, Canada

Monterey Bay Aquarium Research Institute, Moss Landing CA, USA

Dalhousie University, Halifax, Canada

The Natural History Museum, London, UK

University of British Columbia, Vancouver, Canada

University of Geneva, Geneva, Switzerland

Scientific & Technical Advisory Board (STAB)

Linda Amaral Zettler

Claire Gachon

Laure Guillou

Line Le Gall

Laura Wegener Parfrey

Stéphane Audic

Matthew Brown

Micah Dunthorn

Enrique Lara

Frédéric Mahé

Ramon Massana

Alexandros Stamatakis

Woods Hole Marine Biological Laboratory, Woods Hole MA, USA

Scottish Association for Marine Science, Oban, UK

CNRS - Station Biologique de Roscoff, Roscoff, France

Museum National d'Histoire Naturelle, Paris, France

University of British Columbia, Vancouver, Canada

CNRS - Station Biologique de Roscoff, Roscoff, France

Mississippi State University, Mississippi State MS, USA

University of Kaiserslautern, Kaiserslautern, Germany

University of Neuchâtel, Neuchâtel, Switzerland

CIRAD, Montpellier, France

Institut de Ciències del Mar, Barcelona, Spain

Heidelberg Institute for Theoretical Studies, Heidelberg, Germany