# Deliverable D7.2

**Funding Scheme: THEME [ICT-2007.8.0] [FET Open]**

# Paving the Way for Future Emerging DNA-based Technologies: Computer-Aided Design and Manufacturing of DNA libraries

**Grant Agreement number: 265505**

**Project acronym: CADMAD**

**Deliverable number: D7.2**

**Deliverable name: Agenda and Minutes 12 Month Meeting**

| |
|---|
| **Contractual Date[1] of Delivery to the CEC: M12** |
| **Actual Date of Delivery to the CEC: M14** |
| **Author(s)[2]: Gila Yagur** |
| **Participant(s)[3]: OSM** |
| **Work Package: WP7** |
| **Security[4]: Pub** |
| **Nature[5]: R** |
| **Version[6]: 0.1** |
| Total number of pages: |

1 As specified in Annex I

2 i.e. name of the person(s) responsible for the preparation of the document

3 Short name of partner(s) responsible for the deliverable

4 The Technical Annex of the project provides a list of deliverables to be submitted, with the following classification level:
**Pub -** Public document; No restrictions on access; may be given freely to any interested party or published openly on the web, provided the author and source are mentioned and the content is not altered.
**Rest -** Restricted circulation list (including Commission Project Officer). This circulation list will be designated in agreement with the source project. May not be given to persons or bodies not listed.
**Int -** Internal circulation within project (and Commission Project Officer). The deliverable cannot be disclosed to any third party outside the project.

5 **R (Report):** the deliverables consists in a document reporting the results of interest.
**P (Prototype):** the deliverable is actually consisting in a physical prototype**,** whose location and functionalities are described in the submitted document (however, the actual deliverable must be available for inspection and/or audit in the indicated place)
**D (Demonstrator):** the deliverable is a software program, a device or a physical set-up aimed to demonstrate a concept and described in the submitted document (however, the actual deliverable must be available for inspection and/or audit in the indicated place)
**O** (**Other):** the deliverable described in the submitted document can not be classified as one of the above (e.g. specification, tools, tests, etc.)

6 Two digits separated by a dot:
The first digit is 0 for draft, 1 for project approved document, 2 or more for further revisions (e.g. in case of non acceptance by the Commission) requiring explicit approval by the project itself;
The second digit is a number indicating minor changes to the document not requiring an explicit approval by the project.

## Abstract

The 12 Month Meeting of the project CADMAD took place at the Weizmann Institute of Science on April 15 and 16, 2012. It consisted of individual Working meetings of each Work Package on the first day and review of the work done over the first year and future plans for the next 6 months on the second day.


**CADMAD 12 Month Meeting**

April 15&16, 2012
The Weizmann Institute of Science, Rehovot, Israel

# Agenda

| **Day I**<br>***Sunday April 15, 2012***<br>***Location: Room 1 in Ziskind Building*** | | | |
|---|---|---|---|
| ***Time*** | ***Title*** | ***Responsible person*** | ***Location*** |
| 12:00 | Registration in the lobby of Ziskind Building Building<br>Welcome reception – in the Faculty Lounge, ground floor | Ehud Shapiro And staff | Zisking Building |
| **Work Packages specific workgroups**<br>• *Each Work Package participant should attend his relevant WP workgroup.*<br>• *The aim of these workgroups is to define WP technical objectives, description of the work, deliverables and milestones for the next 6 Months, implementation plan and anticipated deliverables*<br>• Each session will start with a presentation by the WP leader which will then go on to lead the session<br>• All Workgroups will take place in Room 261, 2nd floor.<br>An additional room has been reserved for informal meetings from 12:00 until 18:00 (room 141A – in the Faculty Lounge) | | | |
| 12:45 – 13:45 | **WP1** Developing textual and graphical tools for computer-aided DNA library specification | UNOTT | Room 261, 2nd floor |
| 13:45– 14:45 | **WP2** Developing biochemistry and algorithms for a computer-aided DNA design based on DNA reuse | WEIZMANN | Room 261, 2nd floor |
| 14:45 – 15:00 | **Break** | | |
| 15:00 – 15:00 | **WP3** Automation of DNA processing based on DNA reuse | RUB | Room 261, 2nd floor |
| 16:00 – 16:00 | **WP4** Multi-layer system integration and the development of faults detection, isolation and correction methodologies | WEIZMANN | Room 261, 2nd floor |
| 17:00 – 18:00 | **WP5** End users' applications: Directing system development and potency validation | UKB | Room 261, 2nd floor |
| 18:30 | Meet in the lobby of San Martin Guesthouse<br>for dinner in Cezar restaurant at the Science Park in Rehovot | | |

| Day II<br>*Monday April 16, 2012*<br>*Location: Room #1 in Ziskind Building* | | | |
|---|---|---|---|
| ***Time*** | ***Title*** | ***Responsible person*** | ***Location*** |
| From 7:30 a.m. | Breakfast for all the participants<br>served in the small dining room of the San Martin Guesthouse | | San Martin Guesthouse |
| 9:00 | Opening of the Meeting | Ehud Shapiro - WEIZMANN | Ziskind Building Room #1 |
| **Format of this day presentations should fit the format requested for the Review**<br>Topics to be covered in blue | | | |
| 9:00 – 9:25 | Coordinator General Overview on the project<br>- project objectives, team and responsibilities<br>- specific objectives planned for the period to be reviewed<br>- overview of achieved objectives (in terms of deliverables and milestones)<br>- deviations from original plans | WEIZMANN – Prof. Udi Shapiro | Room #1, Ziskind Building |
| ***Work Packages Review of the work done over the first year***<br>- planned objectives (generally and for period under review)<br>- achieved objectives (in terms of deliverables and milestones)<br>- **deviations from original plans** in terms of technical work as well as use of resources<br>- future work and anticipated deliverables | | | |
| 9:25 – 10:00 | **WP1** Developing textual and graphical tools for computer-aided DNA library specification | UNOTT – Prof. Natalio Krasnogor | Room #1 |
| 10:00 – 11:00 | Faculty Seminar given by<br>**Prof. John McCaskill,** Ruhr-University-Bochum<br>Biomolecular Information Processing (BioMIP)<br>Bioorganische Chemie<br>Topic: *"Chemical feedback microprocessors for constructive DNA processing"* | | |
| 11:00 – 11:15 | **Break** | | |
| 11:15 – 12:00 | **WP2** Developing biochemistry and algorithms for a computer-aided DNA design based on DNA reuse | WEIZMANN – Dr. Tuval Ben-Yehezkel | Room #1 |
| 12:00 | **Lunch break** | | In Assia Garden |
| 13:00 – 13: 40 | **WP3** Automation of DNA processing based on DNA reuse | RUB – Prof. John McCaskill | Room #1 |
| 13:40 – 14:20 | **WP4** Multi-layer system integration and the development of faults detection, isolation and correction methodologies | WEIZMANN – Dr. Tuval Ben- | Room #1 |

| | | | |
|---|---|---|---|
| | | Yehezkel | |
| 14:20 – 15:00 | **WP5** End users' applications: Directing system development and potency validation | UKB – Dr. Sandra Meyer | Room #1 |
| 15:00 – 15:15 | **Break** | | |
| 15:15 – 15:25 | **WP6** Dissemination and Exploitation | OSM - Pnina Dan | Room #1 |
| 15:25 – 15:45 | **WP7** Management and Technical Coordination<br>• Review project administrative , financial and reporting procedures<br>• Website management | Pnina Dan<br>Tuval Ben-Yehezkel | Room #1 |
| 15:45 – 16:45 | Discussion, conclusions | All Partners<br>Moderator: Udi/Tuval | Room #1 |
| 16:45 – 17:00 | Fixing next meeting location and date<br>Closing the Meeting | Ehud Shapiro | Room #1 |
| 18:30 | Meeting in the lobby of the San Martin Guesthouse<br>For dinner in Oro restaurant, Rehovot Science Park | Tuval | |

## List of Participants

*

| Acronym | Affiliation | Name | e-mail | Phones |
|---|---|---|---|---|
| Advisor | Agilent Lab. , Israel | Dr. Zohar Yakhini | zohar_yakhini@agilent.com | Tel. +972 4 829 4937<br>Mob. +972 52 4299 728 |
| Advisor | Interlab Ltd., Merck Group, Israel | Dr. Simone Botti | simone.botti@merckgroup.com | Tel. +972 8 9382388<br>Mob. +972 54 6727878 |
| Advisor | Tecan Group, Switzerland | Dr. Marc Feiglin | Marc.Feiglin@tecan.com | Tel. +1 732 698 0363 |
| WEIZMANN<br>P1 | Weizmann Institute of Science, Israel | Prof. Ehud Shapiro | ehud.shapiro@weizmann.ac.il | Tel. +972 544-929-108 |
| | | Dr. Tuval B Yehezkel | tuval.ben-yehezkel@weizmann.ac.il | Tel. +972 528681034 |
| **UNOTT**<br>P3 | University of Nottingham, U.K. | Prof. Natalio Krasnogor | Natalio.Krasnogor@Nottingham.ac.uk | Tel. +44 7825753885 |
| | | Dr. Jonathan Blakes | jvb@cs.nott.ac.uk | Tel. +44 7855259355 |
| UKB<br>P4 | Universitaetsklinikum University of Bonn, Germany | Dr. Sandra Meyer | meyer-sandra@gmx.de | Tel. +49-228-6885-529 |
| RUB<br>P5 | Ruhr-University-Bochum, Germany | Prof. John McCaskill | john.mccaskill@biomip.rub.de | Tel. +49-(0)234-32-27702 |
| | | Antonio Gabriel Minero | Gabriel.Minero@ruhr-uni-bochum.de | |
| ETHZ | Eidgenössische | Prof. Joerg Stelling | joerg.stelling@bsse.ethz.ch | Tel. +41-61-38-73194 |

| | | | | |
|---|---|---|---|---|
| P6 | Technische Hochschule Zürich, Switzerland | Ellis Whitehead | ellis.whitehead@gmail.com | Tel. +41-61-38-73194 |
| | | Prof. Sven Panke | sven.panke@bsse.ethz.ch | Tel. +41-61-387 3209 |
| FMI P7 | Friedrich Miescher Institute, Switzerland | Dr. Arnaud Krebs | arnaud.krebs@fmi.ch | Tel. +41 (0) 616976680 |
| UH P8 | University of Helsinki, Finland | Dr. Konstantin Kogan | Konstantin Kogan <konstantin.kogan@helsinki.fi> | Tel. +358 9 191 58921 |
| OSM P9 | OSM-Dan Ltd., Israel | Dr. Pnina Dan | pninadan@osmdan.com | **Tel. +972 8 9460012 Mob. +972 544567024** |
| | | Gila Yagur | gila@osmdan.com | **Tel. +972 8 9460012 Mob. +972 525563163** |

# Internal work package discussions and plan for future work

| WP1 | Developing textual and graphical tools for computer-aided DNA library specification |
|---|---|
| WP Leader | UNOTT |
| Participants | All |

## Main Objectives

The main objectives of the WP1 discussions were to:

- present work done since the last meeting
- explain some issues we had identified, how these were addressed and their impact going forward
- demonstrate refinements to the DNALD language
- discussion proposed additions to language aimed at improving extensibility and usability
- plan for next 6 months

## Discussion – Summary

Language is not "executable" but "a specification". However, it still needs to be "evaluated"/ "computed" to derive the set of the targets from the expressions that define them, and the components of those targets need to be "traced" for visualisation purposes. The order of evaluation **is** important: the right way is a topologically sorted directed acyclic graph. **Syntactically valid libraries with cyclic dependencies between definitions constitute an error on the part of the library designer.**

DNALD and its Integrated Development Environment (IDE) should be regularly benchmark for usability and utility by our biological partners. To support this (new versions of/ new) libraries will collected and organized into a test harness. **It should be the responsibility of each WP5 partner to keep their library up to date as the language changes**. Additional libraries from WEIZ demonstrated in WP2 discussion should also be included.

Software development needs clear checkpoints to evaluate its progress. The current implementation plan (in order of development priority) is: minimum functionality as set out in deliverables 1.1 and 1.2, improved visualization of libraries, documentation incorporating library examples as above.

We need to know concretely which additional features are most important to users (of which WP5 partners are a representative sample) in order to effectively manage the evolution of the language and IDE towards version 1.0. An example is interfacing sequence databases with the language as sources for library inputs. Arnaud Krebs (FMI) made some suggestions of curated gene/protein DBs: ENSEMBL (http://www.ensembl.org/), RefSeq (http://www.ncbi.nlm.nih.gov/RefSeq/) and reference genomes hosted at UCSC, and helpfully pointed out that Bioconductor's Biostrings.getSeq() function can be used to obtained these. Such leads will be followed up and we (UNOTT) will review the respective APIs and assess feasibility of developing a generic interface[7] for sequence retrieval functions in the language to allow for incremental addition of support for various databases.

As more WP5 partners will have other sources and it is important that we prioritize the most requested, **it was agreed that WP5 leader should coordinate with partners to gather this information**. At the same time other useful information to collect would be:

- which organisms (and strains) groups work with, in order to establish which codon tables are most used

- which third-party sequence-oriented programs partners use, for what purposes and what they like/dislike about the interface

- sketches/mock ups of possible library/sequence visualisations
- capacity for working with libraries, i.e. how many different sequences could be used experimentally per man month

For example UKB works only with E. coli, mouse and human cells, could utilize multiple cross-translation between these and suggested that available codon tables be presented as a filterable list with most common organisms at its head.

Miscellaneous IDE features which came up in the discussion included:

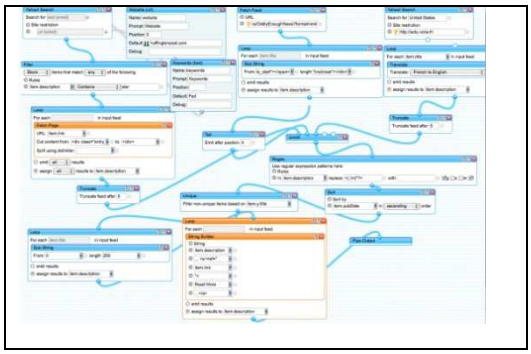
- Library statistics view showing: # output sequences, # sequence fragments, etc. of all libraries in workspace
- Separation of inputs and outputs in visualization (with QSplitter|SWT ExpandBar widget)
- Partial visualization of non-terminals (subexpressions) for debugging.
- Highlighting subsequences with biological functions, restriction enzyme recognition sites (RES) being the most prominent example. In particular, the highlighting of RES unwittingly created by mutations or the concatenation of two sequences in a library design, which are known only once the DNALD file has been interpreted to compute the actual library specification. Differential highlighting of: (actual/potential) secondary sequences or lack thereof, methylated and glycosylated bases, natural/synthetic sequences were all mentioned. The conclusion is that there is a need for a generic means by which users can specify **any** pattern to be highlighted in the editor or visualization, using regular expressions, something simpler like wildcards, ambiguous nucleotides or even DNALD expressions.

Miscellaneous language features which came up in the discussion included:

- Circular inputs and outputs, e.g.: `'ACGT(AC)'`, should be permitted (needed by UNOTT and relevant to plasmids)
- Protection of input sequences (i.e. ability to lock/make readonly)
- Annotations for metadata (e.g. `@genomic @ORFs(7..77,88..888) input1 := 'atagctagctagctag…')`
- Wrappers for (subsets of) outputs which could encode library delivery format (e.g. `5'plasmid ( x := …; y := …) 3'plasmid`)
- Ability to specify DNALD libraries based **entirely** on amino acid sequences (with codon usage tables), rather than on nucleotides. Can be interpreted as either giving the manufacturer maximum freedom over DNA sequences (any backtranslation), or by annotating with a coding sequence, avoiding backtranslation by working with a known DNA sequence but at the amino acid level.
- An alanine scanning function that returns the set of amino acid sequences where each sequence has one amino acid replaced by an alanine)

Increasing complexity of the language, even though driven by real user requirements, is a threat to its stability and coherence. Nottingham has proposed (and this was accepted by all partners) that functionality could be devolved to functions implemented outside of the interpreter and integrated into the IDE itself. Example functions and "extension methods" were presented.

A general mechanism for producing and handling sets of sequences as *generators* was proposed and **accepted**, whereby a generator *lazily* produces a stream of sequences (from the potentially exponential set of sequence it encodes) which are *filtered* to conform to the library designs biological requirements as encoded by the filters. Sources of generators are back-translations (which can be biased according to codon usage), alternatives (don't care non-determinism/degrees of freedom), disambiguations of sequences containing ambiguous nucleotides and de Bruijn sequences.



The generators and filters approach also offers a potential solution to the visualization of DNALD operations, akin to Yahoo Pipes (right):

Interfacing WP1 frontend (library specifications computed by DNA Library Designer) with WP4 backend (the library construction planning algorithm) was another point of discussion. Generally, the frontend should communicate to the backend: ***what*** is required (preserving degrees of freedom[8]). Backend to frontend: which targets are not possible to produce with the given degrees of freedom (and therefore how the library design might be altered to make it feasible); estimate of library costings (time and money).

Currently the planning algorithm receives the sets of input and target sequences defined by the computed specification and attempts to derive a set of the longest common substrings from which to begin planning library construction (called the preprocessing stage). While it may not the case that the library is designed suboptimally, this approach which aims for maximal reuse and provides very a clean interface, is so Spartan that it disregards a large amount of useful information in the library specification (the users intentions). A better approach would be to communicate inputs (with sources) and synthetic sequences (including mutations) as the set of initial *fragments* and each target as a series of concatenations of subsequences of those fragments. Each individual concatenation is a potential intermediate node in the planning tree. **The means by which this structured information should be communicated between the frontend and backend should be further discussed in the next 6 jointly with learder of WP4 (WIS)**

The following points are gathered from the WP1 round table presentation and discussion with all CADMAD partners at the start of the M12 meeting, and in separate discussions between partners in WP1 (UNOTT: Jon, Natalio), WP2 (WEIZ: Tuval), WP3 (ETH: Ellis), WP4 (WEIZ: Ofir) and WP5 (UKB: Sandra, FMI: Arnaud, UH: Constantine)

## Conclusions

- UNOTT to have quarterly meetings with WEIZ (WP4) and UKB (WP5) groups to discussion integration of library specification with planning and end user contributions to language/interface/visualization design.
- Create a versioned repository of libraries for testing each new software release. WP5 partners to update their libraries in accordance with changes to the language with only descriptive assistance from UNOTT. This will form the basis for an anonymised and scrambled set of language examples which demonstrate a variety of DNALD library structures.

- Frontend should provide backend with the set of initial *fragments* and each target as a series of concatenations of subsequences of those fragments, as opposed to simply sequences.
- To further discuss, with WP4(WIS) where computationally expensive (e.g. backtranslation) operations reside (frontend VS backend). For example, since degrees of freedom must be preserved (when possible) between frontend and backend some computationally expensive operations (principally back-translation) that currently is part of the frontend may better be done at the backend and hence not be performed by the frontend. Together with WP4 (WIS) we need to investigate whether it might be better to let the frontend be responsible only for parameterisation of sequence generators (and accompanying filters) as fragments which will be evaluated in the backend (WP4).

✓ To use a rapid application development strategy in which every 3 months new releases (bugs fixed, new features) are delivered to users.

| **Action items** | **Person responsible** | **Deadline** |
|---|---|---|
| ✓ Create private repository with existing WP5 libraries and other from past WEIZ synthesis projects for software testing and tutorials. | Jonathan Blakes | June 2012 |
| ✓ Sharing code for generator data structure and DNA sequence ambiguation algorithms with WEIZ. | Jonathan Blakes | August 2012 |
| ✓ Release new version of DNA Library Designer, with robust evaluation scheme but no new language features, to partners for testing. | Jonathan Blakes | July 2012 |
| ✓ In parallel with their testing we will add read/write support for SBOL and ApE formats, connectivity to one or two sequence databases, improved visualization with the possibility of zooming in/out of the graphical representation of the libraries with labeled sequence fragments (as in the WIS web-based original version), to be delivered incrementally through point releases. | Jonathan Blakes | November 2012 |
| ✓ Prototyping "pipes/flows" visual interface, secondary structure detection/filtering, functional annotations, more databases connections, read/writing GENOCAD and GeneDesigner formats, and other WP5 requested features. Time permitting and automatic updates functionality | Jonathan Blakes | 2nd annual review in 2013 |

**Assigned internal deliverables**

| From | | To | | Deliverable description | Purpose | Due Date |
|---|---|---|---|---|---|---|
| WP4 | WEIZ | WP1 | UNOTT | Original DNApl language interpreter implementation (archive of Yair Mazor's files) | Identify undocumented language features and disambiguate intentions of ++ operator | April 2012 |
| WP2 | WEIZ | WP1 | UNOTT | DNApl files for intronome and translation libraries | Enlarge set of libraries for testing software | April 2012 |
| WP1 | UNOTT | WP4 | WEIZ | Code for sequence generators (and their sampling in the case of backtranslation) and DNA sequence ambiguation | Defer potentially expensive operations to backend | August 2012 |
| WP5 | UKB | WP1 | UNOTT | Lists of sequence DBs, software, organisms and UI sketches provided by partners and prioritized by number of partners that provide them | Obtain and prioritize features to be developed | July 2012 |
| WP1 | UNOTT | WP5 | all | Rolling software releases | Incremental development and feedback | July 2012 onwards |

**Foreseen Internal Meetings**

| Participants | | Purpose | Date | Place |
|---|---|---|---|---|
| Name | Institute | | | |
| **Jonathan Blakes** | UNOTT | Integration of computed DNA library specifications (WP1) with construction planning algorithm inputs (WP5) and communication of planning failures to IDE. | July 2012 | Nottingham, UK or Weizmann Institute, Israel |
| **Ofir Raz** | WEIZ | | | |
| **Jonathan Blakes** | UNOTT | Discuss WP5 partners additional library requirements (language features), graphical user interface storyboards (IDE features for deliverable 1.3) and tutoring in use of the software (to guide suitable end-user documentation). | July | Bonn, Germany or Nottingham, UK |
| **Sandra Meyer** | UKB | | | |

**Special Notes**:

The next planned deliverable for WP1 is a report on GUI features due in month 18 (July 2012).

| WP2 | Developing biochemistry and algorithms for a computer-aided DNA design based on DNA reuse |
|---|---|
| WP Leader | Weizmann |
| Participants | UNOTT, ETHZ |

## Main Objectives

To discuss and plan future work regarding emerging topics related to biochemistry and algorithms for a computer aided DNA design based on DNA reuse.

## Discussion - Summary

Possible modes of cooperation were discussed and initial plans were devised in the context of developing and Integrating several advanced features into the planning phase. Specifically, the development of new primer Software and the development of an interface between the specification and planning phases in the context of Shared fragment analysis was discussed and initial plans for implementation were suggested.
The current planning algorithm prototype successfully handles the planning of explicit targets while input DNA and ordering automatically designed synthetic oligonucleotides.
We plan to enable the design and manufacturing of degenerate targets.
To do so, the DNALD offers many "generators" that add degrees of freedom to the explicit sequence. For Sequences can be designed while incorporating wildcard bases such as 'N' that will "compile" to a set of all explicit sequences that matches the designed pattern.
UNNOT have been developing the designing UI that enables said generation of freedom degrees and the planning algorithm to handle degenerate targets is now our main goal.

## Conclusions

- ✓ Shared fragment analysis should be integrated into the planning algorithms
- ✓ Possible information sharing between vDNApl and the planning algorithms regarding shared fragments in Library should be explored
- ✓ There is a need to develop ways by which the planning algorithm can accommodate degenerate bases in specification of the library.
- ✓ Continue adaptation of reaction conditions to higher throughput using the robotic system
- ✓ Better primer design computational tools should be developed to support the Y operation
- ✓ The communication between the DNALD interface and the planning algorithm will consist of functional and structures instead of explicit target sequences.
- ✓ Those functional objects should incorporate target sequences degeneracy when required.
- ✓ The planning algorithm should support the manufacturing of degenerate targets.

| Action items | Person responsible | Deadline |
|---|---|---|
| Commence primer design project | Tuval | |
| Develop a shared fragment interface | Blakes & Raz | |
| Devise reaction conditions for HT synthesis | Tuval | |
| Extend planning algorithm with degeneracy support | Ofir Raz | |
| Design and implement the design->planning interface | Blakes & Raz | |

**Assigned internal deliverables**

| From | | To | | Deliverable description | Purpose | Due Date |
|---|---|---|---|---|---|---|
| WP1 | UNNOT | WP2 | Weizmann | Library description Interface | Integration | |
| WP2 | Weizmann | WP1 | UNNOT | Library validation | Integration | |
| | | | | | | |

**Foreseen Internal Meetings**

| Participants | | Purpose | Date | Place |
|---|---|---|---|---|
| Name | Institute | | | |
| Jonathan Blakes | UNNOT | Defining the design->planning interface. | August | Nottingham |
| Ofir Raz | Weizmann | | | |
| | | | | |

| WP3 | Automation of DNA processing based on DNA reuse |
|---|---|
| WP Leader | RUB |
| Participants | RUB, UKB, ETHZ, UEVE, FMI, UH |

## Main Objectives

The automation of DNA processing involves a development for liquid handling robots and for microfluidics. The overarching goal of the internal work package discussions was to interrelate these two efforts in WP3, and in particular to explore the extent to which software, which is being developed for the robotic platform, could also interface to microfluidic systems and software development there, and for which type of microfluidic system this would be most straightforward. Finally, the planning objectives also included the new T3.5, on customization of automation for DNA library synthesis, due to start at M19 and the role of RoboEase software customization in that.

## Discussion - Summary

T3.1 and T3.4 In terms of the overall requirements and specifications (T3.1) the role of isothermal amplification protocols as an alternative to PCR were discussed. Despite work at both Weizmann and RUB, characterizing the products of the nSDA reactions for the test system (length 450), the reaction is not yet ready for automated deployment, with concerns over reproducibility and product identity. It was agreed that product sequencing would be performed at Weizmann to verify operation in microfluidics at RUB of amplification reactions.
RUB highlighted significant progress in the novel approach to feedback integrated electrophoretic length separation, but this is also not yet ready for automated deployment. So parallel to the integrated microfluidic strategy developed by RUB, an alternative involving millifluidics and electrowetting, with on chip PCR and magnetic bead based separation, was worked out in a discussion between Tecan advisor (M. Feiglin) and Weizmann and RUB. A trial involving the investigation of a test module for the Y-operation, employing existing dimensions and technology at Tecan (San Jose) on the 2 µl scale, was proposed in a joint visit of RUB and Weizmann to California. This plan still needs confirmation by Tecan company. Overall timescale for approval and test: two months. Additional discussion involved droplet processing in microfluidics and open-planar digital millifluidics. These were compared and contrasted in their abilities to support high density amplification and separation of DNA.

T3.4.5 The microfluidics compiler was discussed, including the IP issues that arose with UEVE and the deliverable 3.7. The agreement was reached that RUB would extend its initial work in developing a compiler (via its RUB-IML) to deliver a solution which goes significantly beyond integrating the interfaces to microfluidics control peripherals, and take over some financial resources and responsibility in this regard from UEVE to achieve this.

Discussions on RoboEase development (led by ETHZ) centered around high-level language capabilities (with backward compatibility), portability (e.g., via common structure for data and commands), and design for debugging (including automatic documentation). Apart from biologist users, integration of RoboEase into the production / automation platform, as well as potentially for control of microfluidics devices will be evaluated.

*T3.4.5 Customizing automated robotics for selected applications that test the DNA libraries (S. Panke, ETHZ)*
After producing DNA libraries (WP5) using the CADMAD platform, these libraries need to be functionally analyzed using user specific downstream assays. Since the size of a typical DNA library to be produced by CADMAD will be tens or hundreds of molecules, such assays should be conducted in a high-throughput manner. Often it is the case that the development of such assays takes between weeks and months to establish. The RoboEase programming language will be used to quickly setup library-specific assays for each library application within WP5. Such assays

may include, for example, cloning of the DNA molecules into bacteria and screening them for a specific property (say fluorescence emission). The selected strains can then be automatically picked, sequenced and can even be selected to be used as input for another DNA processing iteration. Selected applications from WP5 libraries will be automated thus dramatically shorten the time to apply the libraries produced by CADMAD.

**Conclusions**

- ✓ Good progress in microfluidic integration, despite difficulties in adapting biochemistry
- ✓ Test digital droplet millifluidics using Tecan EWOD approach in next 2-3 months
- ✓ Continue work with next generation electronic microfluidics, including microfluidics compiler
- ✓ Development of RoboEase is on track according to plan (first deliverables M24)
- ✓ Collect information on robotic platforms available at partners
- ✓ Months 19 – 24: Implement prototype biochemical assays at ETHZ
- ✓ Transfer adapted RoboEase to partners around months 24
- ✓ Partners set up assays and communicate problems.

| Action items | Person responsible | Deadline |
|---|---|---|
| Test of EWOD millifluidics at Tecan lab in San Jose | McCaskill, Tuval b. E. | 31.7.2012 (Month 18) |
| Collect information on robotic platforms | Panke | 31.8.2012 (Month 19) |
| Transfer Roboease for biochem assay to partners | Panke | 28.2.2013 (Month 25) |
| Transfer of RoboEase programmer documentation from ETHZ to RUB (+1month) | Stelling | 31.5.2012 (Month 16) |
| ETHZ to RUB | | |

**Assigned internal deliverables**

| From | | To | | Deliverable description | Purpose | Due Date |
|---|---|---|---|---|---|---|
| WP# | Partner | WP# | Partner | | | |
| 3 | RUB | 3 | Weizmann | Amplification samples | Sequencing test of products | Several times over next 6 months |
| 3 | Weizmann | 3 | RUB | Sequencing results | Sequencing test of products | Several times over next 6 months |
| 3 | ETHZ | 3 | UKB | Roboease file | Implement assay on robot | 28.2.2013 |
| | | | FMI | See above | See above | See above |
| | | | UEVE | See above | See above | See above |
| | | | UH | See above | See above | See above |
| | | | | | | |
| | | | | | | |

**Foreseen Internal Meetings**

| Participants | | Purpose | Date | Place |
|---|---|---|---|---|
| Name | Institute | | | |
| Ellis Whitehead | ETHZ | Database work | August 2012 | Rehovot |
| Ofir Raz | WEIZ | Production System integration | December 2012 | Basel |
| Tuval, Shapiro McCaskill | WEIZ, RUB | Test of EWOD millifluidics for Y operation | May-June 2012 | San Jose |
| | | | | |

| WP4 | Multi-layer system integration and the development of faults detection, isolation and correction methodologies |
|---|---|
| WP Leader | Weizmann |
| Participants | |

## Main Objectives

**Integrating the various system components.**
Integrate The Roboease language developed at ETHZ with the live manufacturing processes, back-end database model and automation module.
Integrate the DNALD library developed at UNNOT as input.
Couple the back-end database model with the manufacturing process.
Adapt the planning algorithm's output to the new database model.
Improve QC measures,

## Discussion - Summary

The main link between the various system components is the database model. Well defined manners of Input and output should be defined and adapted to the data model.

Live database connectivity will enable a unified data structure for the entire project and dynamic feedback to the various interfaces involved.

The Roboease language will incorporate a production database module. That database model should become one with the system's database model.

The Roboease language will be "aware" of its reagents database representation and will incorporate such data into it's logic. (e.g. resolve locations and available volumes of reagents, update that data after consumption of Reagent).

## Conclusions

- ✓ All system's interfaces should be DB-bound.
- ✓ Integration should include live database connectivity to enable the previous feature.
- ✓ Explore methods for automated GE analysis

| Action items | Person responsible | Deadline |
|---|---|---|
| Integrate DNALD with the planning module | Blakes&Raz | |
| Integrate planning module's I/O with the database | Ofir Raz | |
| Integrate automation module's I/O with the database | Ofir Raz | |
| Integrate Roboease's I/O with the database | Whitehead&Raz | |
| Implement DB-bound lab-work protocols | Ofir Raz | |

**Assigned internal deliverables**

| From | | To | | Deliverable description | Purpose | Due Date |
|---|---|---|---|---|---|---|
| WP4 | Weizmann | WP3 | ETHZ | Reagents database model | Handling production reagents | |
| | | | | | | |

**Foreseen Internal Meetings**

| Participants | | Purpose | Date | Place |
|---|---|---|---|---|
| Name | Institute | | | |
| Jonathan Blakes | UNNOT | Integrating WP1 by defining the design->planning interface. | August | Nottingham |
| Ofir Raz | Weizmann | | | |
| Ellis Whitehead | ETHZ | Integrating the Database model with the Roboease WP | August | Basel |
| Ofir Raz | Weizmann | | | |
| | | | | |

| WP5 | End users' applications: Directing system development and potency validation |
|---|---|
| WP Leader | UKB |
| Participants | WEIZMANN, UEVE, UNOTT, ETHZ, FMI, UH |

## Main Objectives

The main objectives of this work package are (1) to focus and direct the development of the CADMAD platform to current and future DNA programming requirements, (2) to validate the produced libraries and (3) to compare the CADMAD libraries against libraries made by existing technologies. There was one deliverable in the first year of the project, namely the report on defining user specifications and requirements from DNApl and vDNApl (GUI) (D5.3). To fulfill this deliverable the end users designed eight DNA libraries using the newly developed DNA library designer (DNald). Based on the particular necessities of these libraries further specific requirements for the DNA programming language were defined.

## Discussion - Summary

The compiled list of requirements and specifications for the DNA programming language was presented and the necessity for further advancement of the programming language was discussed. The priority in which the defined requirements should be implemented in the software was established. It was suggested that the end users should provide a list of already existing programmes that feature the desired tasks, so that the used algorithms may be integrated into the programme.

The possibility to implement further functions as plug-ins into the software was discussed and agreed upon. It should be discriminated between mandatory and optional plug-ins/functions, which are not necessary for the design of a wide variety of DNA libraries but only for a limited number of them.

It was decided to first re-design the present DNA libraries using an improved version of DNald and provide feedback to UNOTT and then start to design more complex DNA libraries to further test and challenge the DNA programming language.

## Conclusions

- ✓ Re-design present DNA libraries with current version of DNald and provide feedback to UNOTT/UKB
- ✓ Provide a list of already existing programmes that feature desired applications to UNOTT/UKB
- ✓ Re-design present DNA libraries with advanced version of DNald which already includes some of the defined requirements and specifications and provide feedback to UNOTT/UKB
- ✓ Design more complex libraries using advanced version of DNald

| Action items | Person responsible | Deadline |
|---|---|---|
| First re-design of present DNA libraries | WP5 members | May 2012 |
| Search for already existing programmes that feature desired applications and send to UKB | WP5 members | May 2012 |
| Send a compiled list of programmes/desired applications to UNOTT | UKB | June 2012 |
| Second re-design of present DNA libraries | WP5 members | Aug 2012 |
| Finish end user's more complex library drafts | WP5 members | Dec 2012 |

**Assigned internal deliverables**

| From | | To | | Deliverable description | Purpose | Due Date |
|---|---|---|---|---|---|---|
| WP5 | All | WP 5 | UKB | Re-design of present libraries using advanced version of DNald and presentation of visual output | Test and challenge advanced version of DNald and define further end users' requirements | Sep 2012 |
| WP5 | UKB | WP 1 | UNOTT | Re-design of present libraries using advanced version of DNald and presentation of visual output | Test and challenge advanced version of DNald and define further end users' requirements | Oct 2012 |
| WP5 | All | WP 5 | UKB | High level description of application libraries | Demonstrate that CADMAD is able to address multiple parameters of diverse end users' needs | Dec 2012 |
| WP5 | UKB | WP 1 | UNOTT | High level description of application libraries | Demonstrate that CADMAD is able to address multiple parameters of diverse end users' needs | Jan 2013 |

**Foreseen Internal Meetings**

| Participants | | Purpose | Date | Place |
|---|---|---|---|---|
| UKB | UEVE | Discussion about definition of requirements and specifications for DNApl and vDNApl | 13th Feb 2012 | UEVE |
| UKB | UNOTT | Discussion about implementation of further functions in DNald and providing feedback concerning the usability of the programming language | July 2012 | UNOTT or or UKB |
| FMI | UKB | Discussion about DNA libraries and definition of further requirements and specifications for DNApl and vDNApl | Sep/Oct 2012 | UKB |

**Special Notes**:
The following list of requirements and specifications was used as the basis of the discussion:

1. Integration of a **graphical interface** to give the possibility to "drag-and-drop" parts of combinatorial DNA
2. Possibility to **zoom-in and zoom-out** in the visual presentation to facilitate "drag-and-drop"
3. Highlight **restriction enzyme sites** in the output section to facilitate subsequent cloning
4. Exclude **toxic sequences** (e.g. unwanted restriction sites, cryptic splice sites)
5. Integration of a possibility to filter and select against **DNA secondary structures**
6. Inclusion of a function to define the **reverse complement** of a sequence without the need to include the reverse sequence in the input section
7. Integration of a function for **backtranslation**
8. Integrate **codon optimization tables** for expression in various host cells
9. Integrate **DNA optimizer algorithms** such as Gene Optimizer® (Raab et al., 2010)
10. **Compatibility** with open source DNA construction software (e.g. ApE)

11. **Delivery format** (e.g. 96-well-plate)and **form of library output** (e.g. linear DNA fragment) should be selectable
12. **Platform-independent** DNA programming language
13. **Quality control** of the output DNA has to be assured

| WP6 | Dissemination, Exploitation, training and Education |
|---|---|
| WP Leader | OSM |
| Participants | All Partners |

## Discussion - Summary

Dissemination activities were discussed to include:

1. Publishing in peer review journals
2. Participation in conferences
3. Website
4. Workshop – Internal
5. B.Sc. and M. Sc. Courses

Publishing in peer reviews journals: 7 articles were published
Participation in conferences: 19 publications in conferences throughout Europe
Website: the project website was established and is maintained by WEIZMANN

The first year internal workshop: Prof. John McCaskill, Ruhr-University-Bochum, Biomolecular Information

Processing (BioMIP) gave an Invited lecture at the 12M Meeting of CADMAD project. The lecture was given within the Faculty of Mathematics and Computer Sciences on the topic: *"Chemical feedback microprocessors for constructive DNA processing". Faculty scientists and students, and all CADMAD partners attended this lecture.*

B.Sc. and M. Sc. Courses: Sven Panke and Joerg Stelling of ETHZ gave a course on Synthetic Biology including DNA assembly. It consisted of 14 lectures, 3 hours each, starting 21.2.11 in front of 52 students at ETHZ.

### Conclusions

Dissemination activities should be encouraged: Teresa de Martino recommends CADMAD team to target also wide dissemination media channels. As an example she gave a recent article on DNA Computing http://www.economist.com/node/21548488 which is reporting mainly on American on-going research (not yet results!!!). She hopes to see CADMAD soon on some highly visible media.

| WP7 | Management and Technical Coordination |
|---|---|
| WP Leader | OSM |
| Participants | WEIZMANN |


## Summary

The following items were covered:

1. Management Structure
2. Contractual provisions
    - Deliverables
    - Milestones
    - First Periodic Report
3. GA Amendment
4. Financial

Project Officer and reviewers were introduced:

- Project Officer, Dr. Teresa de Martino European Commission, Project Officer European Commission
- Reviewer - Prof. Annie Viallat, University of Marseilles, Laboratoire Adhesion & Inflammation, CNRS,
- Reviewer - Prof. Gioia Altobelli , Queen Mary University of London Centre for Endocrinology, William Harvey Research Institute
- Reviewer – Dr. Anna Battisti, Interdisciplinary Lab. For Computational Science, Bruno Kesler Foundation, Trento, Italy

The Advisor Board was introduced:

- Dr. Zohar Yakhini - Agilent Co.& Technion, Israel
- Dr. Simone Botti - Interlab Ltd., Merckx Group, Israel
- Dr. Marc Feiglin - Tecan Group , Switzerland

Responsibilities of the Advisory Board:

- Review the overall technological and scientific strategy of the project
- Generate a technical assessment of the completed and running activities
- Advice mitigation actions or other upon request

## Internal meetings and Deliverables

- Twenty Internal deliverables were exchanged between the partners
- Four Internal meetings

## Contractual Provisions: Deliverables and Milestones

The status of the first period deliverables was reported:

14 deliverables were submitted for Period 1. Exception:

D3.7 There are IP questions about this deliverable thus withheld and will be delivered during P2.

D6.3 Internal Workshop – deliverable will be submitted after the workshop

D7.2 Agenda and Minutes 12 Month Meeting - Will be delivered after the meeting

Milestone:

MS4 – completion of Robo-Ease v1.0 – WP3 RUB – M18

**P1 Periodic report**

The P1 Periodic Report was prepared and submitted to the NEF and sent to the project officer and the reviewers before the Review Meeting.

**GA Amendment**

UEVE withdrawal
Requested documents:

1. Official amendment request
2. Letter from the coordinator to justify UEVE withdrawal
3. Letter from UEVE that accepts the withdrawal
4. Updated DoW and GPF

**Financial**

First part of the advance payment was transferred to the partners.
Second part will be transferred within the next few days.

Next 35 % was transferred by the coordinator since all requested documents for the first 12M were submitted by the partners to the coordinator and OSM.