LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-739169

Phase II: Automated System for Aneuploidy Detection in Sperm Final Report CRADA No. TC-1554-98

W. J. Wyrobek, R. T. Dunlay

September 29, 2017

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Phase II: Automated System for Aneuploidy Detection in Sperm

Final Report
CRADA No. TC-1554-98
Date Technical Work Ended: February 28, 2003

Date: May 1, 2003 Revision: 4

A. Parties

This project was a relationship between Lawrence Livermore National Laboratory (LLNL) and Cellomics, Inc.

The Regents of the University of California
Lawrence Livermore National Laboratory
7000 East Avenue
Livermore, CA 94550
Andrew J. Wyrobek, LLNL Principal Investigator
Tel: (925) 422-6296
Fax: (925) 424-3130

Cellomics, Inc.
100 Technology Drive
Pittsburgh, PA 15219
R. Terry Dunlay, Cellomics Principal Investigator
Tel: (412) 770-2270
Fax: (412) 770-2201

B. Project Scope

This was a collaborative effort between the University of California, Lawrence Livermore National Laboratory (LLNL) and Cellomics, Inc. (formerly BioDx and Biological Detection, Inc.) to develop an automated system for detecting human sperm aneuploidy. Aneuploidy (an abnormal number of chromosomes) is one of the major categories of chromosomally abnormal sperm, which results in chromosomally defective pregnancies and babies. An automated system would be used for testing the effects of toxic agents and for other research and clinical applications. This collaborated effort was funded by a National Institutes of Environmental Health Services, Phase II, Small Business Innovation Research Program (SBIR) grant to Cellomics (Contract No. N44-ES-82004).

Major deliverables in Phase II included semi-annual progress reports, a Phase II Final Report and a set of software programs for the automated system for detection of DNA probes in sperm. Progress reports were delivered on a semi-annual basis, and several extensions were needed.

However, a commercially ready turn-key system capable of detecting aneuploidy in human sperm was not achieved, but additional system requirements were identified. The final software program provided by the company was evaluated, and is suitable for the collection of "truth data" of sex chromosome probes that would facilitate the development of classifiers for the future detection of sex-chromosomal aneuploidy in sperm.

The following tasks and deliverables were planned to be completed under this project.

Phase II – Planned Cellomics TASKS:

1) Improve image analysis algorithms to accurately detect sperm nuclei and probes labeled by multicolor FISH. The desire is to develop algorithms to detect with > ~98% accuracy aneuploidy in sperm with a very low frequency of chromosomal abnormalities (~10 out of 10,000 cells). (months 1-24)

    a) Develop software System Evaluation and Training (SET) tools to enable a human scorer to interactively assess and record the results of the computer-based segmentation. The SET tool will allow a large number of measurements to be stored along with the biologist's expert judgment as to whether a signal is "probe" or "non-probe" (or "nucleus" vs. "non-nucleus") so that algorithms can be trained with this "ground truth" or "gold standard" training set. (months 1-12)

    b) Use training data collected by LLNL using the SET tool to develop improved algorithms to accurately identify single sperm nuclei and to provide a segmentation mask that accurately represents the nucleus. (months 18-24)

    c) Use training data collected by LLNL using the SET tool to develop improved algorithms to assure accurate detection of probes located in multiple focal planes in human sperm nuclei and to provide a segmentation mask that accurately represents the probe. (months 18-24)

2) Develop "rare event" classification schemes for detecting specific categories of aneuploid sperm among a large excess of normal cells. (months 18-24)

    a) Incorporate the developed inference calculus into the automated scanning system to provide a "ranking" algorithm that grades cells according to probability of abnormality. (months 18-24)

3) Validate the system for the detection of aneuploid human sperm. (months 1-24)

    a) Develop software to enable a human scorer to interactively score and validate the system on specific categories of rare aneuploid cells. (months 1-6)

Phase II – Planned LLNL Tasks

1) Develop "rare event" classification schemes for detecting specific categories of aneuploid sperm cells among a large number of normal sperm cells. (months 1-18)

   a) Use the SlideScan system demonstrated in Phase I to obtain large numbers of images representing various types of aneuploid and normal sperm. It is expected to be necessary to survey several 10,000's cells for this task and will include the preparation, abstracting, and transmission of large amounts of data. (months 1-18)

   b) Use the SET software to enable a human scorer to interactively collect data on specific categories of rare aneuploid cells as well as normals. (months 12-18)

   c) The resulting "training sets" will be used by LLNL to train a Binary Decision Tree classifier. The BDT will exploit all possible feature relationships and determine which data interactions are best suited to the automatic discrimination of probes. (months 12-18)

   d) LLNL will develop inference calculus strategies for combining the confidence probabilities for each detected probe so that each nucleus can be ranked in order of probability of being aneuploid. (months 18-24)

2) Validate the system for the detection of aneuploid human sperm. (months 12-24)

   a) Perform manual scoring to determine if the proposed system works. Use the SET tool as a new extended capability of the automated sperm aneuploidy detection system to quantitatively validate the accuracy of the system using a large number of test slides. (months 12-24)

   b) Conduct side-by-side comparisons of the frequencies of aneuploid sperm determined visually by the biologist and by the trained system. (months 18-24)

3) Use SET tool to perform system validation of at least one probe from running the system in automatic scanning mode on several test slides with known aneuploidy levels.

Phase II – Planned Joint Tasks

1) Engineer from Cellomics will work on site at LLNL with LLNL biologist and engineer at LLNL to:

   a) Develop the "rare event" classification scheme and validate the system for detection of aneuploid human sperm. (months 1-21)

   b) Optimize the hardware for configuration of the aneuploidy detection system. (months 1-21)

2. Use results from the LLNL-generated Binary Decision Tree to build a classifier for probes and nuclei. This classifier will report a confidence level for each probe and nucleus. (months 18-24)

3. Work with LLNL to develop an inference calculus, which will incorporate the confidence levels for all, probes in a nucleus and compute the probability that the nucleus is aneuploid. (months 18-24)

4) Perform an in-depth validation of the system for detection of aneuploid human sperm for at least one probe. (months 18-24)

5) Optimize hardware and software to improve system throughput. (months 1-18)

   a) Determine optimal combination of microscope objective and camera resolution. (months 1-12)

   b) Optimize camera operation for faster autofocus. (months 1-6)

   c) Optimize system software to improve throughput. (months 12-18)

Deliverables

- Establish and complete high-speed Internet link connection, install new PC, microscope and camera hardware at Cellomics; calibrate and test systems by scanning slides. (Cellomics)
- Install PC and relocate system in laboratory; calibrate and test system. (LLNL)
- Develop System Evaluation and Training (SET) tool. (Cellomics)
- Prepare test slides. (LLNL)
- Complete development of nuclei segmentation algorithms. (Cellomics)
- Using SET tool, train Binary Decision Tree (BDT) classifier. (LLNL)
- Develop enhanced SET tool for system validation experiments. (Cellomics)
- Perform final system evaluation of three probes using automatic scanning mode. (LLNL)
- Phase II final report. (LLNL and Cellomics)

C. Technical Accomplishments

Cellomics Phase II Technical Accomplishments

- Established a high-speed T1 internet link between LLNL and Cellomics
- Developed an experimental plan for test images to be used for writing, fine tuning and testing of the image segmentation routines and the probe spot quantitation algorithms.

- Designed a software infrastructure consisting of a large library of image segmentation routines, a number of image object measurement routines and the Visilog image processing package.
- Upgraded from Windows 3.1 to Windows NT to utilize multiprocessing and multithreading capabilities
- Utilized advanced object-oriented methodology to simplify the software and algorithm development process
- Enhanced image collection capabilities, including flexible scanning algorithms, better auto-focus and auto-exposure algorithms, and a highly sophisticated data storage component
- Added extensive data and image visualization tools
- Developed Image Acquisition tool
- Developed System Evaluation and Training (SET) tool
- An illumination calibration tool was created
- New camera drivers were developed
- The nuclei segmentation and chromosome detection algorithms were completed for normal sperm cells
- A complete beta version of the system was installed at LLNL
- Improved speed by revamping the autofocusing algorithm, optimizing the stage speed and reorganizing the scanning algorithm
- Improved the Calibration Tool
- Upgraded the automation controller
- Modified SlideScan to enable storage of images acquired at multiple focal depths
- Modified the assay algorithm
- Substantially redesigned the SET Tool
- Developed a program to generate reports from data generated using the SET Tool

LLNL and Joint Technical Accomplishments

- Evaluated microscope magnification, probe signal, and focal fields, and identified optimal microscope and image capture configurations
- Evaluated sequential versions of the Cellomics software for the detection of sperm nuclei, and probes used in the aneuploidy assays, and made recommendations to improve the software functionality

- Evaluated the final version of the Cellomics software for the accuracy of detection of X, Y and 21 probes, and established the high accuracy of the detection of X and Y probes in normal human sperm cells.
- Evaluated that the final version of the Cellomics software and found an accuracy of ~65% for chromosome 21 probe among normal cells, which was insufficiently low for the detection of aneuploidy in sperm.
- Evaluated the final version of the Cellomics software and described additional system requirements for the accurate detection of all three probes used in the aneuploidy assay.

D. Expected Economic Impact

D.1 Specific Benefits

No immediate economic impact is expected from the results of this project. The successful completion of this project would have benefited several research groups in the Biology and Biotechnology Research Program at LLNL, as well as germ cell research and fertility medicine at an international level. It might also have provided the basis for an efficient machine system to score chromosomal defects in somatic cells. Such a system would have ready applications in basic research, epidemiology, toxicology, etc., in the general research community, and would be specifically relevant for the new DOE missions of genetic susceptibility and functional genomics. Unfortunately, a fully functional automated system for detecting chromosomally damaged cells was not produced. The existing product, at the completion of Phase II, is capable of collecting data for a training set that would aid in creating the classifiers necessary for the future research needed to develop an automated system to detect chromosomally abnormal sperm cells.

E. Partner Contribution

Major project contributions by Cellomics, Inc. are summarized in Section C. above. While several of the deliverables were made that allowed the detection of sperm nuclei under the microscope as well as the accurate automated detection of two of the three probes of the aneuploid system, the goal of developing a ready to commercialize system for the detection abnormal sperm cells was not achieved.

F. Documents/Reference List

Reports

Automated Image Analysis system for Human Sperm: A proposal submitted to FY94 DOE/DP SMALL BUSINESS INITIATIVE PROGRAM. MAY 4, 1994

Automated System for Aneuploidy Detection in Sperm: SBIR Phase I proposal submitted to NIEHS. December 4, 1995.

Automated System for Aneuploidy Detection in Sperm: Small Business Innovation Research Phase I Final Report. April 30, 1997.

Automated System for Aneuploidy Detection in Sperm: SBIR Phase II proposal submitted to NIEHS. June 6, 1997.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: April 8, 1998 through July 8, 1998.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: July 8, 1998 through January 30, 1999.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: February 1, 1999 through July 15, 1999.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: July 16, 1999 through December 31, 1999.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: January 1, 2000 through June 30, 2000.

NIEHS Small Business Innovative Research (SBIR) Topic No. 48 Phase II
"Automated System for Aneuploidy Detection in Sperm"
Reporting period: July 1, 2000 through December 31, 2000.

Copyright Activity
Cellomics developed several software programs under this CRADA.

Subject Inventions
No subject inventions resulted from this CRADA.

Background Intellectual Property

LLNL had disclosed the following Background Intellectual Property for this project:

Software:

LLNL Code Release No. 960005 - *Sperm Morphometry Automation Program* (MAP), V.5.0, Authors: Andrew J. Wyrobek, Michael Firpo.

Invention Disclosures:

IL-7866: Patent pending

IL-8467B: Patent pending

We do not currently anticipate any licensing of the LLNL BIP.

G. Acknowledgement

Participant's signature of the final report indicates the following:

1) The Participant has reviewed the final report and concurs with the statements made therein.
2) The Participant agrees that any modifications or changes from the initial proposal were discussed and agreed to during the term of the project.
3) The Participant certifies that all reports either completed or in process are listed and all subject inventions and the associated intellectual property protection measures generated by his/her respective company and attributable to the project have been disclosed and included in Section E or are included on a list attached to this report.
4) The Participant certifies that if tangible personal property was exchanged during the agreement, all has either been returned to the initial custodian or transferred permanently.
5) The Participant certifies that proprietary information has been returned or destroyed by LLNL.

R. Terry Dunlay, Principal Investigator — Date
Cellomics, Inc.

Andrew J. Wyrobek, Principal Investigator — 8/15/03 Date
Lawrence Livermore National Laboratory

Karena D. McKinley, IPAC Director — 8/20/03 Date
Lawrence Livermore National Laboratory

Attachment I – Final Abstract

Phase II: Automated System for Aneuploidy Detection in Sperm

Final Abstract (Attachment I)
CRADA No. TC-1554-98
Date Technical Work Ended: February 28, 2003

Date: May 1, 2003 Revision: 4

A. Parties

This project was a relationship between Lawrence Livermore National Laboratory (LLNL) and Cellomics, Inc.

The Regents of the University of California
Lawrence Livermore National Laboratory
7000 East Avenue
Livermore, CA 94550
Andrew J. Wyrobek, LLNL Principal Investigator
Tel: (925) 422-6296
Fax: (925) 424-3130

Cellomics, Inc.
100 Technology Drive
Pittsburgh, PA 15219
R. Terry Dunlay, Cellomics, Inc. Principal Investigator
Tel: (412) 770-2270
Fax: (412) 770-2201

B. Purpose and Description

This was a collaborative effort between the University of California, Lawrence Livermore National Laboratory (LLNL) and Cellomics, Inc. (formerly BioDx and Biological Detection, Inc.) to develop an automated system for detecting human sperm aneuploidy. Aneuploidy (an abnormal number of chromosomes) is one of the major categories of chromosomally abnormal sperm which results in chromosomally defective pregnancies and babies. An automated system would be used for testing the effects of toxic agents and for other research and clinical applications. This collaborated effort was funded by a National Institutes of Environmental Health Services, Phase II, Small Business Innovation Research Program (SBIR) grant to Cellomics (Contract No. N44-ES-82004).

Major deliverables included in phase II included semi-annual progress reports, a Phase II Final report and a set of software programs for the automated system for detection of DNA probes in sperm. Progress reports were delivered on a semi-annual basis, and several extensions were

needed. However, a system capable of detecting aneuploidy in human sperm was not achieved, but additional system requirements were identified. The final software program provided by the company was evaluated, and is suitable for the collection of "truth data" that would facilitate the future development of classifiers for the detection of aneuploidy.

C. Benefit to Industry

Recently several FISH-based assays have been developed to detect a range of chromosomal abnormalities in human and rodent sperm. These assays facilitate the understanding of the mechanisms of such abnormalities that lead to birth defects and inherited diseases. However, extensive use of such assays has been limited by the lack of an automated scoring system that would increase the speed of analysis and save many man-hours.

Epidemiological evidence shows that environmental or occupational exposure of fathers to toxic agents can lead to abnormal reproductive outcomes including spontaneous abortions, birth defects and childhood cancer. One possible mechanism for this is the transmission of genetic damage via sperm. Aneuploidy is one of the major types of genetic defects in sperm. In animal studies, numerical aneuploidy in male germ cells can arise from abnormalities in the meiotic spindle, centromeres, etc. which have been induced by agents such as colchicine, vinblastine, chloral hydrate, and possibly alcohol. In humans, aneuploidy at birth predominantly includes chromosomes 13, 18, 21, X, and Y, and the paternal contribution is most significant for sex chromosomal aneuploidies.

The analysis of sperm-FISH assays for chromosomal abnormalities is very labor intensive. Fluorescent signals have to be scored in thousands of cells while constantly adjusting the plane of focus to detect fluorescent signals in different focal planes in the three-dimensional sperm head. It takes an experienced microscope technician 10-15 hours to score one slide of 10,000 cells. For these reasons, an automated system is critical for making the recently developed sperm-FISH assays part of routine screening for infertility and for widespread applications in genetic toxicology.

D. Benefit To DOE/LLNL

The successful completion of this project would have benefited several research groups in the Biology and Biotechnology Research Program at LLNL. It would have provided an efficient machine system to score chromosomal defects in germ cells and somatic cells. Such a system would have ready applications in basic research, epidemiology, toxicology, etc., in the general research community, and would be specifically relevant for the new DOE missions of genetic susceptibility and functional genomics. Unfortunately, a fully functional automated system for detecting chromosomally damaged cells was not produced. The existing product is capable of collecting data for a training set that would aid in creating the classifiers necessary for an automated system for the detection of chromosomally abnormal sperm cells.

E. Project Dates

We executed the project timeline tabulated below through the sixth six-month period and then suspended work on the project. Plans for a seventh and eighth periods were cast based upon the prospect of follow-on funding to complete the work.

First Report Period (April – June 1998) - DONE:

New Task #	*Description*
0.1	Hire technical personnel.
0.2	Order equipment.
0.3	Write first semi-annual status report for the start-up period.

Second Report Period (July – December 1998) - DONE:

New Task #	*Description*
1.1	Prepare test slides at LLNL and acquire image data for algorithm development.
1.2	Development and implementation of the SET tool.
1.3	Continue development of nuclei segmentation algorithms.
1.4	Begin developing probe quantization algorithms for first pass candidate cell detection.
1.5	Begin developing multi-focal-plane probe segmentation algorithms for second pass probe enumeration measurement routines.
1.6	Start developing Graphical User Interface (GUI) for scanning application and SET tool.
1.7	Write second semi-annual status report for this period.

Third Report Period (January – June 1999) - DONE:

New Task #	*Description*
2.1	Continue development and implementation of the SET tool.
2.2	Continue development of nuclei segmentation algorithms.
2.3	Continue developing probe quantization algorithms for first-pass candidate cell detection.
2.4	Continue developing multi-focal-plane probe segmentation algorithms for second pass probe enumeration measurement routines.
2.5	Continue developing Graphical User Interface (GUI) for scanning application and SET tool.
2.6	Write third semi-annual status report for this period.

Fourth Report Period (July – December 1999) - DONE:

New Task #	***Description***
3.1	Complete development of nuclei segmentation algorithms; evaluate algorithms.
3.2	Start evaluating classification algorithms using preliminary image data.
3.3	Complete development of the SET tool.
3.4	Prepare slides for extensive learning and test data sets; acquire image data.
3.5	Finish framework, GUI and image processing components of the software.
3.6	Complete development of tool to calibrate illumination source.
3.7	Assemble the complete beta version of the software.
3.8	Port the software to the LLNL system.
3.9	Write fourth project period semi-annual report.

Fifth Report Period (January – June 2000) - DONE:

New Task #	*Description*
4.1	Investigate hardware system optimizations.
4.2	Write fifth semi-annual report.

Sixth Report Period (July – December 2000): DONE

New Task #	***Description***
5.1	Determine the microscope objective to be employed and the number of image planes required.
5.2	Complete development of software required for truth data collection.
5.3	Start testing and refining of systems at LLNL.
5.4	Design classifier evaluation and training facilities.
5.5	Write semi-annual report for this project period.

Seventh and Eight Report Periods (January – December 2001*):

New Task #	*Description*
6.1	LLNL will evaluate the modified SET Tool for efficiency of detecting nuclei as well as the probes for chromosomes X, Y, and 21. LLNL will train a student and collect truth data for normal cells and the specified categories of aneuploid cells.
6.2	Finish the algorithms for feature extraction.
6.3	Develop classifier evaluation and training facilities.
6.4	Evaluate algorithms for classification and ranking.
6.5	Integrate classification algorithms into AutoFISH system.
6.6	Deliver incremental improvements in classifier and GUI components to LLNL for evaluation.
6.7	LLNL will use SET tool to optimize the sample preparation for automatic analysis.
6.8	Finish implementing all optimizations and deliver the version 1.0 system.
6.9	Write Phase II final report.

* These plans were made contingent upon receipt of continuation funding. We elected not to pursue this funding. Funding was suspended and work was terminated. Work was restarted in fall of 2002 and the evaluation of the last version of the Cellomics software was evaluated for probe detection accuracy. The X and Y probes were detected a high accuracy of ~95% or greater, but the autosomal probe was detected at only ~60% accuracy. Additional system requirements were defined accordingly in the final phase II report delivered in March 2003.