DOE/ER/61233-7

RECEIVED
SEP 06 2000
OSTI

DEVELOPMENT AND APPLICATION OF SUBTRACTIVE HYBRIDIZATION
STRATEGIES TO FACILITATE GENE DISCOVERY

Prepared for

THE U.S. DEPARTMENT OF ENERGY
AGREEMENT No. DE-FG02-91ER61233

Marcelo Bento Soares (PI) and Greg Lennon (Co-PI)

Columbia University
722W 168th St, Box#41, New York, N.Y. 10032

August 1997

We have no objection from a patent
standpoint to the publication or
dissemination of this material.

Nathaniel Dvorsky Aug 30 2000
Office of Intellectual
Property Counsel
DOE Field Office, Chicago
Date

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed or represents that its use would not infringe privately-owned rights.

ABSTRACT

We have developed and applied subtractive hybridization strategies to generate cDNA libraries enriched for novel cDNAs. Pools of I.M.A.G.E. clones, from which ESTs have been derived, are used as drivers in hybridizations with single or multiple normalized libraries thus generating subtracted libraries enriched for cDNAs not yet represented in public databases. Subtracted libraries are characterized by Southern hybridization to assess reduction in representation of clones of the driver population and then contributed to the I.M.A.G.E. consortium for large-scale arraying and sequencing.

1. INTRODUCTION

The methods we originally developed to normalize directionally cloned cDNA libraries (Soares et al., 1994; Bonaldo, Lennon & Soares, 1996) have been successfully utilized to generate a number of human, mouse and rat cDNA libraries. All human and mouse libraries (and soon the rat libraries as well) have been contributed to the I.M.A.G.E. consortium and they have been extensively used for large scale generation of expressed sequence tags (ESTs). Both the ESTs and their respective clones are publicly available.

Although the use of normalized libraries has proven most advantageous to minimize the redundant identification of the mRNAs of the super-prevalent and intermediate frequency classes within a particular tissue, it cannot prevent the redundant identification of mRNAs (of any frequency class) that are expressed in multiple tissues. In other words, normalization alone cannot avoid the redundant identification of ESTs that have been obtained previously from other libraries. This problem is becoming increasingly more relevant as we approach completion of the ongoing human and mouse gene discovery efforts. Hence, we proposed to take advantage of subtractive hybridization strategies that we developed, to generate libraries enriched for novel cDNAs. It is anticipated that the use of subtracted libraries will become increasingly advantageous as we strive towards the ultimate goal of identifying all human and mouse genes. This project has two major goals: (1) to optimize the method for construction of subtracted libraries, and (2) to generate subtracted libraries to facilitate the ongoing human and mouse EST programs. Briefly, we have further optimized our method and we have applied it to generate a second subtracted library. Just as it was the case for the first subtracted human fetal liver-spleen library that we generated, this second subtracted library proved to be significantly enriched for clones not represented in the driver population, thus facilitating the identification of novel cDNAs. Following is a description of the method that we used, and of the results that have been obtained.

2. CHARACTERIZATION OF A MELANOCYTE-FETAL HEART-PREGNANT UTERUS- SUBTRACTED LIBRARY

We have documented the feasibility of two approaches to generate subtracted libraries enriched for novel cDNAs: (1) within libraries, and (2) across libraries. In the first approach, subtraction within libraries, a pool of I.M.A.G.E. clones derived from a single library are used as a driver in a hybridization with the normalized library from which those clones were originally derived. We have applied this strategy to generate a subtracted human fetal liver-spleen library. A total of 5,000 I.M.A.G.E. clones originally derived from the normalized fetal liver-spleen library were used as a driver in a hybridization with the

normalized fetal liver-spleen library itself. Characterization of the subtracted liver-spleen library by Southern hybridization with a number of cDNA probes derived from the driver population clearly indicated that the procedure had been successful to reduce the representation of the driver population in the subtracted library. The subtracted library was then submitted for arraying and large-scale sequencing. The results of the analysis of the EST data generated thus far firmly documented that the procedure was successful to reduce the representation of the driver population by about 4 to 5 fold.

Once we had documented the feasibility of the first strategy, we proceeded to assess the feasibility of the second approach, i.e. subtraction across libraries. For that we used as driver a pool of 15,000 I.M.A.G.E. clones that had been derived from (5,000 each) three normalized libraries: human fetal heart, human melanocytes and human pregnant uterus. A mixture of the three normalized libraries (the tracer) was then hybridized with a vast excess of this driver and the subtracted library was tested for the occurrence of clones represented in the driver population. Once again, Southern hybridizations of endonuclease-restricted library DNA with probes derived from the driver population strongly indicated that the procedure was successful. Most importantly, however, sequence analysis of a large number of ESTs derived from this subtracted library firmly demonstrated that the procedure was effective to reduced the representation of the driver population in the subtracted library by about 3.5 to 4.3 fold.

3. PROPOSED PLAN

Now that we have firmly demonstrated the feasibility of both strategies, we will proceed with the original plan of generating human and mouse subtracted libraries enriched for novel cDNAs. In close collaboration with Greg Lennon (Lawrence Livermore National Laboratory) we will use pools of I.M.A.G.E. clones derived from the several libraries in hybridizations with mixtures of our best normalized libraries. The subtracted libraries will be characterized by Southern hybridization using probes derived from the driver population, and once proven to be of adequate quality they will be shipped to Lawrence Livermore National Laboratory for arraying in Greg Lennon's laboratory, and from there to the Genome Sequencing Center at Washington University for generation of ESTs.

4. LITERATURE CITED

- Bonaldo, M. F., Lennon, G. & Soares, M.B. (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Research* 6, 791-806.
- L. Hillier, G. Lennon, M. Becker, M. Bonaldo, B. Chiapelli, S. Chisoe, N. Dietrich, T. DuBuque, A. Favello, W. Gish, M. Hawkins, M. Hultman, T. Kucaba, M. Lacy, M. Le, N. Le, H. Mardis, B. Moore, M. Morris, C. Prange, L. Rifkin, T. Rohlfing, K. Schellenberg, M. Soares, F. Tan, E. Trevaskis, K. Underwood, P. Wohldmann, R. Waterston, R. Wilson and M. Marra. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Research* 6, 807-828.
- Lennon, G.G., Auffray, C., Polymenopoulos, M. & Soares, M.B.(1996). The I.M.A.G.E. Consortium: An integrated molecular analysis of genomes and their expression. *Genomics* 33:151-152.
- Soares, M.B., Bonaldo, M.F., Su, L., Lawton, L. & Efstratiadis, A. (1994). Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. USA* 91(20): 9228-9232.

Inventions

INVEN_NO	CONTRACTNO	SUBCONTRACT	CTRNO	CTRNAME	DOCKETNO	SERIALNO	PATENTNO
91261	DE-FG02-91ER61233		C00027	COLUMBIA UNIVERSITY	42840-B	465,811	5,637,685
81221	DE-FG02-91ER61233		C00027	COLUMBIA UNIVERSITY	352	126,594	5,482,845
92292	DE-FG02-91ER61233		C00027	COLUMBIA UNIVERSITY	42840	619,542	5,830,662
84032	DE-FG02-91ER61233		C00027	COLUMBIA UNIVERSITY	474	465,857	5,702,898
86881	DE-FG02-91ER61233		C00027	COLUMBIA UNIVERSITY	615	715,941	5,846,721

07/24/2000

User: WILLIAMS

IPL DOE CH

	DATE	TIME	TO/FROM	MODE	MIN/SEC	PGS	JOB#	STATUS
16	08/01	09:00	712123053697	EC--S	01'20"	003	123	OK

Fax Transmission**INTELLECTUAL PROPERTY LAW**

9800 South Cass Avenue

Argonne, IL 60439

630-252-2166

Fax: 630-252-2779

To: Marilyn Wallace**Date:** July 25, 2000 8-1-00**Fax #** 212-305-3697**Pages:** 3 including this cover sheet.**From:** Carol Williams**Subject:** Contract No. DE-FG02-91ER61233

This office has the responsibility for conducting a patent review of the subject subcontract and arranging patent clearance for final payment.

I am faxing a copy of the patent certification that was submitted to this office. Five invention disclosures were reported under the subject contract, however, only one is listed on your certification. I have added page 2 and listed the other four. Please review the additions and if you concur, sign and return.

Upon receipt of the patent certification, I will continue with the closeout process of your contract.

Note that the U.S. Government has certain rights in your invention, and the contract under which the subject invention was conceived, provides that your title in the invention may be adversely affected if you do not comply with the terms of the contract.

Thanks.

Please respond.

Fax Transmission
INTELLECTUAL PROPERTY LAW
9800 South Cass Avenue
Argonne, IL 60439
630-252-2166
Fax: 630-252-2779

To: Marilyn Wallace **Date:** July 25, 2000 8-1-00
Fax # 212-305-3697 **Pages:** 3 including this cover sheet.
From: Carol Williams
Subject: Contract No. DE-FG02-91ER61233

This office has the responsibility for conducting a patent review of the subject subcontract and arranging patent clearance for final payment.

I am faxing a copy of the patent certification that was submitted to this office. Five invention disclosures were reported under the subject contract, however, only one is listed on your certification. I have added page 2 and listed the other four. Please review the additions and if you concur, sign and return.

Upon receipt of the patent certification, I will continue with the closeout process of your contract.

Note that the U.S. Government has certain rights in your invention, and the contract under which the subject invention was conceived, provides that your title in the invention may be adversely affected if you do not comply with the terms of the contract.

Thanks.

Please respond.