Designing Microarrays and Microarray Experiments, Statistically, Under Northern Lights

Jason C. Hsu
in collaboration with
Jane Chang, Tao Wang, Youlan Rao, Yoon Lee,
Eiríkur Steingrímsson, Magnús Karl Magnússon, Kristin Bergsteinsdottir

Ah, an invitation to tell stories, even if the audience is not quite captive … (usual FDA-like disclaimer: I apologize in advance for any unintentional insult)

If was dark when I arrived in Iceland, one January morning in 2004. My host (my 2nd PhD student) Gunnar Stefansson took me from the airport to the apartment, and went skiing in the Alps with his family (slight exaggeration). The next day, it snowed. To forage for food, I used a large cast iron skillet to dig myself out of the apartment.

I went to Iceland as a Fulbright Scholar. Some Fulbright Scholars go there to paint, some go there to write. As a mere statistician, I went there to help University of Iceland to develop a Statistics curriculum.

But, secretly, I hoped to connect with scientists working on genomics in Iceland. I had noticed much of the reported findings in bioinformatics do not seem reproducible. Is the promised pot of gold at the end of the –omics rainbow a myth? (see Fig. 1) Iceland is technologically advanced, especially in the genomics area. The company called deCODE Genetics is in the news all the time. Perhaps I can find the answer in Iceland.

Figure 1: Skógarfoss, Iceland
(Legend has it there is a treasure chest behind the waterfall.)

Personally, I believe a major reason for the irreproducibility of bioinformatic results is a lack of basic statistical considerations in the design of such experiments. I have had experience with pharmaceutical clinical trials since 1998. Results from those trials are highly reproducible. There is no doubt a main reason for the reproducibility is these trials are designed and executed according to statistical design principles (randomization, replication, blocking) set forth in the international guidance ICH E9. If one were to follow the same principles in designing microarray experiments, one would randomize the placement of the probes and the samples on the microarrays, take replicate samples from each patient, and hybridize samples from different groups to be compared in blocks. Such design considerations have hardly penetrated the realm of microarray experiments.

Most stock microarrays allow only one biological sample to be placed on each array. With one patient's sample per array, it is impossible to separate array to array variability from patient to patient variability. In comparing expression levels of low risk and high risk patients, the observed differences due to patients belonging to different risk groups are completely confounded with potential differences due to array processing.

Iceland is small: the majority of the Icelandic population lives in Reykjavik, and the majority of the Reykjavik population can comfortably fit into the Ohio Stadium. Small size + High tech = Flexibility, so perhaps such designs are more possible in Iceland. As

an independent academic, I would go to the end of the earth to prove my point. In this case, it was the beginning of the earth I went to: in the novel *Journey to the Center of the Earth* by Jules Vern, the journey starts in Iceland.

With few properly trained biostatistician who can talk multiple testing and pretend to talk –omics in Iceland, I readily connected with scientists at deCODE and a small rival genomic company called Urður, Verðandi, Skuld (UVS). One day, as I was leaving UVS, I noticed it shared its building and laboratory with a company called NimbleGen Iceland! Surfing the web that night, I found NimbleGen to be a maker of microarrays that allow flexible designs. As I walked back to my apartment, under Northern Lights, I thought "This has possibilities." (huge understatement)

Making a *possibility* a *reality* is tough. To prove statistical concepts, ideally one conducts experiments with known answers. For example, one can place titrated RNA samples on microarrays to compare the results from statistical designs versus haphazard designs. Typical funding agency reaction upon receiving such a proposal is why it would fund an experiment with known answers.

Fortunately, we were able to get the Icelandic Government to fund a series of proof-of-concept microarray experiments, designed and analyzed by Ohio State and Bowling Green statisticians, with microarrays synthesized by NimbleGen, and samples prepared and hybridized by UVS.

Our experiments utilized microarrays with mini-microarrays on them, 12 mini-microarrays on each array. Figure 2 demonstrates a Latin Square design to compare high concentration (C_H) with low concentration (C_L) RNA samples from breast cancer cell line (T_b) and colon cancer cell line (T_c). Each microarray serves as a block, with equal numbers of samples from each group to be compared on each microarray. Such a design not only avoids confounding with array effects, it enhances the sensitivity of group comparisons as well. To avoid bias due to position of the probes, the placement of the probes is randomized for each microarray when it is synthesized, separately for each mini-microarray. To avoid bias due to position of the samples, the rows and columns of the Latin Squares are randomized before hybridization.

Figure 2. Example of Latin Square Design of Microarray Experiment

Array 1

T_bC_L	T_bC_H	T_cC_L	T_cC_H
T_bC_H	T_cC_L	T_cC_H	T_bC_L
T_cC_L	T_cC_H	T_bC_L	T_bC_H

Array 3

T_cC_L	T_cC_H	T_bC_L	T_bC_H
T_cC_H	T_bC_L	T_bC_H	T_cC_L
T_bC_L	T_bC_H	T_cC_L	T_cC_H

Array 2

T_cC_H	T_bC_L	T_bC_H	T_cC_L
T_bC_L	T_bC_H	T_cC_L	T_cC_H
T_bC_H	T_cC_L	T_cC_H	T_bC_L

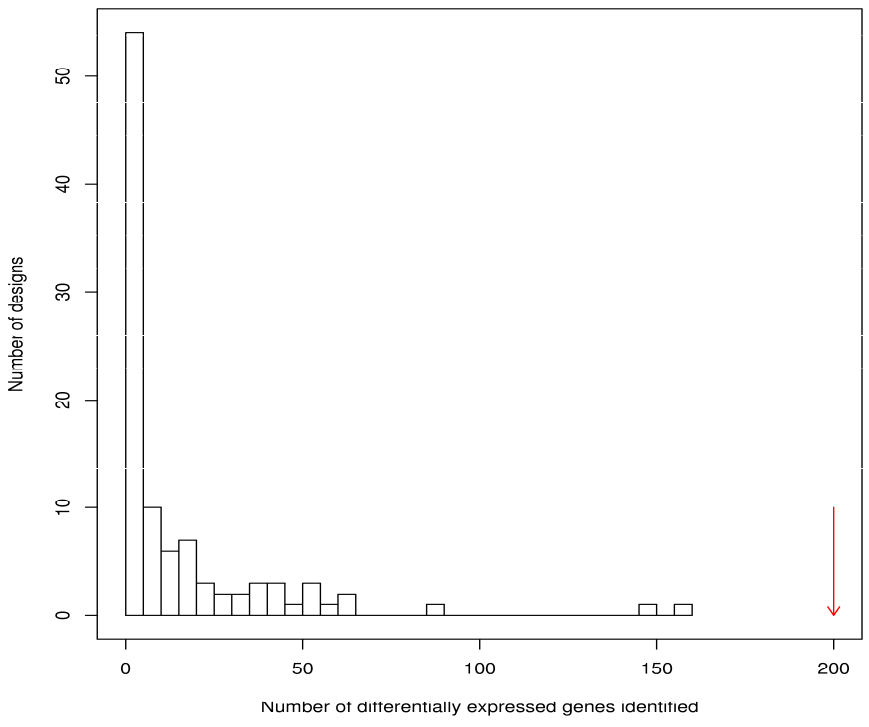
Array 4

T_bC_H	T_cC_L	T_cC_H	T_bC_L
T_cC_L	T_cC_H	T_bC_L	T_bC_H
T_cC_H	T_bC_L	T_bC_H	T_cC_L

To conduct studies that have not been done before is to develop new techniques and software. It is learning to communicate and to appreciate different views. I am grateful to all my collaborators for their efforts and their willingness to risk failure with me.

Result of the first of our experiments was published in Hsu et al (2007). It demonstrates that statistical design of microarrays and microarray experiments can enhance sensitivity and specificity. Figure 3 shows, for the breast cancer cell line, whereas the statistically designed microarray study found all 200 genes to be differentially expressed between high and low concentrations, haphazard designs found fewer genes to be differentially expressed.

Figure 3. Number of genes inferred differentially expressed
breast cancer cell line
Haphazard Design vs. Statistical Design



In the meantime, microarray experiments are moving from for "discovery" only toward "clinical use". In 2005, the FDA issued its Voluntary Pharmacogenomic Data Submission (VGDS) guidance (FDA 2005), which couples the development of a drug for a subgroup of the patient population with the development of a device (e.g., a microarray) that can accurately predict which patients will be responders to the drug. We have conducted a proof-of-concept experiment appropriate for VGDS, which my collaborators and I are in the process of documenting.

The second evening I was in Iceland, I went alone to a Math Department party, in a house next to a large cemetery. Walking past its inhabitants after the party (after schnapps), I thought how fitting a description of that party would be as an ending to my story, should I get to tell it. It is a play on a typical Prairie Home Companion radio broadcast ending:

> "So that's the story from Reykjavik, Iceland, where at math parties all the men (math professors) stay in one room, all the women (wives of professors) stay in another room, and all the children (students) stay in a 3rd room and have a good time."

References

Hsu, Jason C., Chang, Jane, Wang, Tao, Steingrímsson, Eiríkur, Magnússon, Magnús Karl, Bergsteinsdottir, Kristin (2007). Statistically designing microarrays and microarray experiments to enhance sensitivity and specificity. *Briefings in Bioinformatics* **8**(1):22-31. Epub 2006 Aug 9, bb1023. PMID: 16899493

FDA (2005). Pharmacogenomic Data Submission: Guidance for Industry, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration.

ICH E9 (1998). Statistical Principles for Clinical Trials. International Conference on Harmonisation.