Award Number: W81XWH-10-1-0675

TITLE: Genome-Wide Association Study in African-Americans with Systemic Lupus Erythematosus

PRINCIPAL INVESTIGATOR: John Harley, M.D., Ph.D.

CONTRACTING ORGANIZATION: Children’s Hospital Medical Center
Cincinnati, Ohio, 45229-3026

REPORT DATE: September 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
Systemic lupus erythematosus (lupus) is a potentially deadly systemic autoimmune disease that disproportionately afflicts women and African-Americans. This project is designed to discover genes that increase the risk of lupus in African-American. Our goal was to expand the genotyping density by genotyping a subset of our African-American lupus cases and controls on the OMNI-1S platform and then to exploit this genotyping with sequencing and follow up genotyping in an effort to identify the genes that alter disease risk. Two years ago we abandoned our hope of genotyping the control samples from Detroit and instead found and have already genotyped 3000 African-American controls on the OMNI-Express. (These reagents were purchased by our collaborator.) At this point the genotyping is completed and we are working on the quality control and data analysis. As we work through the remaining issues (population stratification, imputation, and detailed exploration of positive findings) we are hopeful to be composing our genome wide association study manuscript in the months to come and then turn our attention to the follow up studies with sequencing data and follow up genotyping. The project now appears to be spectacularly productive with a substantial number of previously unknown genetic associations that are newly detected by this project.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Body</td>
<td>1</td>
</tr>
<tr>
<td>Key Research Accomplishments</td>
<td>1 - 4</td>
</tr>
<tr>
<td>Reportable Outcomes</td>
<td>4</td>
</tr>
<tr>
<td>Conclusion</td>
<td>5</td>
</tr>
<tr>
<td>References</td>
<td>5</td>
</tr>
<tr>
<td>Appendices</td>
<td>None</td>
</tr>
</tbody>
</table>
1. Abstract/Introduction (SF298 requirement)

Systemic lupus erythematosus (lupus) is a potentially deadly systemic autoimmune disease that disproportionately affects women and African-Americans. This project is designed to discover genes that increase the risk of lupus in African-American. Our goal was to expand the genotyping density by genotyping a subset of our African-American lupus cases and controls on the OMNI-1S platform and then to exploit this genotyping with sequencing and follow up genotyping in an effort to identify the genes that alter disease risk. Two years ago we abandoned our hope of genotyping the control samples from Detroit and instead found and have already genotyped 3000 African-American controls on the OMNI-Express. (These reagents were purchased by our collaborator.) At this point the genotyping is completed and we are working on the quality control and data analysis. As we work through the remaining issues (population stratification, imputation, and detailed exploration of positive findings) we are hopeful to be composing our genome wide association study manuscript in the months to come and then turn our attention to the follow up studies with sequencing data and follow up genotyping. The project now appears to be spectacularly productive with a substantial number of previously unknown genetic associations that are newly detected by this project.

2. Body

Since the population history of African ancestry appears to reach back in time much further to the most recent small founder population (~200,000 years) than the other major human ancestries (<50,000 years for Asian, European, or Amerindian), the extent of linkage disequilibrium is much lower in population samples of African ancestry. This means that the usual approach for finding genetic association using haplotype block tagged markers will be less successful in this ancestry. One way to partially compensate for this problem is to increase marker density, which is what we are funded to do in this project in our genetic study of systemic lupus erythematosus (lupus).

Lupus in African-Americans is more severe and more deadly than in other populations, and especially so compared to European-Americans. Indeed, lupus affects women ten times more frequently than men with a strong tendency to strike during the child-bearing years and is relatively common among the Active Duty Military.

3. Key Research Accomplishments

- Our DOD project is a component of a larger project to more fully characterize African-American genetic association with lupus by genome wide association. Genome wide association genotyping is completed. We continue the intensive evaluation of the data, removing markers and samples that appear to have the potential of artifact from stratification, poor clustering, batch effects, reduced marker calling, and disproportions in controls. At this point we are working with data from all sources from 7,300 DNA samples. Of these this project contributed 574 controls and 1,590 cases evaluated on the OMNI-1S platform. These results are complemented by genotyping from 434 controls and 2,359 cases on the OMNI-1 platform. (Together, the OMNI-1S and OMNI-1 are almost equivalent to the OMNI-2.5.) We have 1,536 controls available with genotyping from the OMNI-2.5 platform, available from dbGaP. In addition, we genotyped 3,985 African-American controls on the OMNI-Express platform in a collaboration with Mt. Sinai in New York City. The merged data provide nearly 3 million single nucleotide markers from 7,300 subjects for a dataset that will have the power for inquiry equivalent to nearly 22 billion genotypes when the imputation across the platforms is completed.
• We have discovered presumed artifacts that originate from batch effects between controls sets, which we have eliminated from the data.

• The major issue has been that the imputation of genotypes against the 1000 genomes project was technically difficult to perform and has required a year of diligent effort to solve. As a consequence we have an abundance of positive associations that do not appear to be robust. Consequently, we submitted a request for extension without additional funds in order to address the issues raised by these results, which was submitted to the DoD on 27 August, 2013. We are awaiting a decision on this request. The following four bullets are taken directly from that request.

• The delays in our plans came from the genotyping data having been performed on four different array platforms. This made the imputation of genotyping very difficult, crashing our computer system repeatedly and requiring much sophisticated informatic intervention and experimentation. Indeed, we started the imputation in September 2012 and did not finish the somatic chromosomes until July 2013. The X-chromosome is now being imputed and we hope will be completed before the end of September.

• We cannot order the confirmatory reagents until we are settled on what results need our experimental attention. At this point, we are close to knowing what results need closer
experimental scrutiny, but we are not quite ready, given the still pending X chromosome results. We have saved some funding for this purpose and would intend to do this and any sequencing that we might be able to afford from the remaining funds.

- The outstanding issue has to do with results that are imputed. With a small error accompanies imputation, which is compounded when there is any issue with the genotyping data upon which the imputation is based. Consequently, we approach the imputation results with some reservation that they are correct. Since we have so many apparently positive results we evaluated these data to find the most trustworthy results that we would then intend to publish. Figure 1 presents the initial imputation results for the somatic chromosomes. We imposed additional data quality criteria on these results requiring that there be a 90% call rate, that Hardy Weinberg proportions have p>0.001 in the controls, that the samples done with OMNI-1 controls did not differ from the other controls by p<0.001, that there be a genotyped marker with an association of p>0.001 that is within 20 kb that is in disequilibrium with an $r^2>0.2$. After the loci that do not pass these criteria are removed, we are left with the HLA region, which contains multiple imputed associations, and 100 others, as presented in Figure 2, that have a predicted genome wide imputed association (p<5x10^{-8}). These results are extraordinary and we think it important to determine how many of these are robust, which we would test by direct genotyping in a subset of the subjects.
The critical experiment that we hope to perform with the extension without additional funds (EWAF) is the actual genotyping of as many of the imputed associations as possible in as many subjects as we can afford. Consequently, we will use our remaining funds to perform this critical experiment. Perhaps, we will be able to establish the validity of the association of at least some of these loci (Figure 2) with actual experimental data.

Studies published this past year have presented novel findings in a variety of settings. Molinaros et al (1) shows how applying the strategy of admixture mapping can lead to the detection of genetic effects, in this case showing the complex relationship of IFIH1 to lupus. Vaughn et al (2) reviewed the current state of lupus genetics with an emphasis on B cell signaling. Namjou et al (3) discovered a unique variant of the C1qA gene, known in the null form or when anti-C1qA autoantibodies are present to be associated with lupus. Finally, Kim et al published data showing that the ICAM1-OCAM4-ICAM5 locus is associated with lupus.

We look forward to completing this project and hope that we will have a comparatively productive final report next year.

4. Reportable Outcomes


5. Conclusions

The planned project to expand the density and number of markers in a case-control study of African-American systemic lupus erythematosus is proceeding toward completion.

6. References


7. Appendices

NONE.