

Review

The Lupus Family Registry and Repository

Astrid Rasmussen¹, Sydney Sevier^{1,2}, Jennifer A. Kelly¹, Stuart B. Glenn¹, Teresa Aberle¹, Carisa M. Cooney³, Anya Grether¹, Ellen James¹, Jared Ning¹, Joanne Tesiram¹, Jean Morrisey¹, Tiny Powe¹, Mark Drexel¹, Wes Daniel¹, Bahram Namjou¹, Joshua O. Ojwang¹, Kim L. Nguyen¹, Joshua W. Cavett¹, Jeannie L. Te¹, Judith A. James^{2,4}, R. Hal Scofield^{1,2,5}, Kathy Moser¹, Gary S. Gilkeson⁶, Diane L. Kamen⁶, Craig W. Carson⁷, Ana I. Quintero-del-Rio^{8,9}, Maria del Carmen Ballesteros^{8,9}, Marilyn G. Punaro^{10,11}, David R. Karp¹¹, Daniel J. Wallace¹², Michael Weisman¹³, Joan T. Merrill¹⁴, Roberto Rivera¹⁵, Michelle A. Petri¹⁶, Daniel A. Albert¹⁷, Luis R. Espinoza¹⁸, Tammy O. Utset¹⁹, Timothy S. Shaver²⁰, Eugene Arthur²¹, Juan-Manuel Anaya²², Gail R. Bruner¹ and John B. Harley^{1,2,5}

Abstract

The Lupus Family Registry and Repository (LFRR) was established with the goal of assembling and distributing materials and data from families with one or more living members diagnosed with SLE, in order to address SLE genetics. In the present article, we describe the problems and solutions of the registry design and biometric data gathering; the protocols implemented to guarantee data quality and protection of participant privacy and consent; and the establishment of a local and international network of collaborators. At the same time, we illustrate how the LFRR has enabled progress in lupus genetics research, answering old scientific questions while laying out new challenges in the elucidation of the biologic mechanisms that underlie disease pathogenesis. Trained staff ascertain SLE cases, unaffected family members and population-based controls, proceeding in compliance with the relevant laws and standards; participant consent and privacy are central to the LFRR's effort. Data, DNA, serum, plasma, peripheral blood and transformed B-cell lines are collected and stored, and subject to strict quality control and safety measures. Coded data and materials derived from the registry are available for approved scientific users. The LFRR has contributed to the discovery of most of the 37 genetic associations now known to contribute to lupus through 104 publications. The LFRR contains 2618 lupus cases from 1954 pedigrees that are being studied by 76 approved users and their collaborators. The registry includes difficult to obtain populations, such as multiplex pedigrees, minority patients and affected males, and constitutes the largest collection of lupus pedigrees in the world. The LFRR is a useful resource for the discovery and characterization of genetic associations in SLE.

¹Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, ²Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, ³Division of Plastic and Reconstructive Surgery, Southern Illinois University School of Medicine, Springfield, IL, ⁴Clinical Immunology Research Program, Oklahoma Medical Research Foundation, ⁵U.S. Department of Veterans Affairs Medical Center, Oklahoma City, OK, ⁶Division of Rheumatology and Immunology, Medical University of South Carolina, Charleston, SC, ⁷Oklahoma Arthritis Center, Edmond, OK, ⁸Department of Biochemistry, Division of Genetics, Ponce School of Medicine, Ponce, ⁹Department of Pediatrics, Universidad Central Del Caribe, Bayamon, PR, ¹⁰Pediatric Rheumatology Service, Texas Scottish Rite Hospital, Dallas, ¹¹Rheumatic Diseases Division, University of Texas Southwestern Medical Center, Dallas, TX, ¹²Rheumatology Arthritis Disorders, ¹³Division of Rheumatology, Cedars-Sinai Medical Center, West Hollywood, CA, ¹⁴Clinical Pharmacology Research Program, Oklahoma Medical Research

Foundation, Oklahoma City, OK, ¹⁵Hospital Dr. Susoni, Arecibo, PR, ¹⁶Division of Rheumatology, Department of Medicine, Johns Hopkins University, Baltimore, MD, ¹⁷Division of Rheumatology, Dartmouth-Hitchcock Medical Center, Lebanon, NH, ¹⁸Section of Rheumatology, Louisiana State University Health Sciences Center, School of Medicine, New Orleans, LA, ¹⁹Section of Rheumatology, Pritzker School of Medicine, University of Chicago, Chicago, IL, ²⁰Arthritis and Rheumatology Clinics Kansas, Wichita, KS, ²¹McBride Clinic, Oklahoma City, OK, USA and ²²Center for Autoimmune Diseases Research, School of Medicine and Health Sciences, Universidad del Rosario, Bogotá, Colombia.

Submitted 30 April 2010; accepted 12 August 2010.

Correspondence to: John B. Harley, Cincinnati Children's Hospital Medical Center, Division of Rheumatology, MLC 4010, 3333 Burnet Avenue, Cincinnati, OH 45229, USA. E-mail: john.harley@cchmc.org

Key words: Systemic lupus erythematosus, Registry, Repository, Autoimmune diseases, Genetics, Heritability, Genome-wide association studies, Linkage analysis, Minorities, Women.

Introduction

SLE is a debilitating and potentially fatal autoimmune disease manifested as a wide variety of clinical features and characterized by the production of antibodies to components of the cell. As the immune system in SLE patients is unable to differentiate some specific antigens from healthy tissue and foreign invaders, the consequent inflammatory response affects the joints, skin, blood and many organs, including the kidneys, brain, heart and lungs.

Disease genetics are important, but at this time the biologic mechanisms that link associated genes to disease phenotypes are largely unknown. Genetic associations often introduce us to previously never contemplated biologic possibilities, therefore such genetic variants must be critical components in one or more of the mechanisms that produce the disease phenotype. The Lupus Family Registry and Repository (LFRR), formerly known as the Lupus Multiplex Registry and Repository (LMRR), was established in 1995 with the goal of assembling and distributing materials and data from families with one or more living members diagnosed with lupus in order to address SLE genetics.

This goal is today a reality; the registry and repository have allowed the creation of an interwoven web of efficient and complementary research that has advanced the important work of elucidating the genetics of SLE. There are 76 approved researchers and collaborators that are directly using LFRR resources for their work, 62 of them within the USA and 14 international scientists. Today, the LFRR is operating with 31 full- and part-time staff members who recruit participants; review medical records; staff the laboratory; design, implement and maintain information technology; obtain informed consent from and protect human subjects; analyse the data; and administer the overall effort. Additionally, five affiliated centres across the USA contribute biological samples and clinical information to the registry (details available as supplementary data at *Rheumatology* Online; supplementary data LFRR Approved Scientific User Application Packet has been published with permission from the Oklahoma Medical Research Foundation).

History of the LFRR

SLE is known to occur in families more frequently than in the general population; 5–10% of SLE patients have a second family member with SLE [1] and the sibling recurrence rate (prevalence in siblings of SLE affected/population prevalence) is estimated to be between 8 and 30% [2]. The monozygous twin concordance rate of SLE is 24%, with a lifetime rate concordance of ~30% [3, 4, 5]. Collectively, these observations strongly suggest that an important understanding of the pathogenesis of SLE could

be initiated through the discovery of genes that increase familial susceptibility to the disease.

In the early 1990s, Dr John Harley, at the Oklahoma Medical Research Foundation (OMRF), began assembling multiplex families for SLE, and caught the attention of the National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS) at the American College of Rheumatology (ACR) annual meeting in the fall of 1993. This was the conceptual birthplace of what became the Lupus Multiplex Registry and Repository and led to contractual funding from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) in the fall of 1995.

Establishing a formal registry led to an increased effort in recruitment of multiplex families with SLE, along with more rigorous and detailed data collection and organization, eventually becoming a centralized repository of data, sera, plasma, DNA, transformed B-cell lines and peripheral blood mononuclear cells. The LMRR offered approved scientists access to more than 5 000 000 data points for each sample, including clinical information, demographics, lupus serology and genotyping data from microsatellites, and later, single-nucleotide polymorphisms (SNPs).

Lupus is a disease of disparities: it is estimated that nine out of ten people diagnosed with lupus are women, and four of ten are African-Americans who are at least three times more likely to have lupus than European-Americans of the same sex [6]. Recent data from admixture studies show that American-Indian ancestry is almost 8-fold more important in generating lupus than is European ancestry [7]. Finally, Asian ancestry also leads to a higher prevalence of lupus than European ancestry [8, 9]. Knowing these disparities, the staff and leadership of the LFRR have consistently targeted recruitment efforts towards the population lupus affects the most, maintaining at least 40% minority enrollment since its inception.

A central focus of the repository since 1999 has been identifying the genes responsible for lupus in African-Americans. This effort was boosted by collaboration, in 2003, with Drs Gary Gilkeson and Diane Kamen at the Medical University of South Carolina (MUSC), who spearheaded the enrollment of SLE patients and controls from the Gullah population. The Gullah are a semi-isolated population residing in the Sea Islands along the South Carolina coast and adjacent inland communities. They have lived in this region since the early 1700s when they were transported from West Africa, and constitute a unique population with greater genetic homogeneity than most other African-American communities in the USA. The Gullah maintain stable family units, making assembly of a collection of pedigrees for the study of multi-genetic diseases like lupus possible.

The LFRR genetics approach has evolved over time; originally, it was focused on linkage in multiplex pedigrees and microsatellite genotyping. The technical revolution in SNP genotyping changed the focus to case-control studies for association, making genome-wide association studies (GWASs) feasible. The LFRR responded to this challenge by expanding recruitment to include simplex families, which are ideal for population-based case-control and family-based association studies.

By 2006, the expanded future goals of the LMRR led to active recruitment of both sporadic and familial SLE and the LMRR was renamed the Lupus Family Registry and Repository (LFRR). In addition, this change allowed for a greater variety in race, ethnicity, gender and age of participants diagnosed with SLE, leading to a collection of lupus patients that more closely reflects the natural distribution of the disease across those same variables (Figs 1–3).

Fifteen years after its modest beginnings, the LFRR has collected 2618 affected SLE individuals from 1954 pedigrees, 8157 controls and information on 82 000 individuals and 35 million data points. The bio-bank contains more than 750 000 aliquots of biological samples and the genotyping core has generated more than 650 million genotypes. All of this has been translated into 104 peer-reviewed publications based on access to LFRR resources (see supplementary data available at *Rheumatology* Online) and has made a major contribution toward the transformation of our understanding of the genetics of lupus that, thanks partly to the LFRR, is well under way.

Nowadays, the processes of the LFRR are divided into multiple specialized areas as follows.

Recruiting

The LFRR recruiters receive information, with the permission of the prospective participant, concerning whether they are willing to participate and whether they might satisfy entry criteria. Collaborating investigators, the Centers for Medicare and Medicaid Services, physician referrals, recruitment events, and the LFRR Web site (www.lupus.omrf.org) are ascertainment sources. Each SLE patient who agrees to participate completes a screening interview. Recruiters score the interviews based on the scoring system developed to gauge the ACR classification criteria [10], enter the information into the database, and alert record reviewers of an interview awaiting evaluation. Recruiters proceed based on the record reviewer's decision to enrol the participant, to place her/him into pending, or to remove from further consideration.

Enrolment involves obtaining informed consent from the participant and participating family members and controls, providing paperwork and arranging phlebotomies. Weekly follow-up through phone calls, e-mails or letters is critical to ensure sample and data return. Participants who are hesitant about completing participation after enrolment are removed, but not before substantial recruiter effort to facilitate their initial decision to participate.

Fig. 1 Types of pedigrees enrolled in LFRR. Multiplex: two or more family members affected with SLE participating; Familial: one participant affected with SLE who has or had family members with SLE, but the other SLE affected family members are currently unavailable for participation OR a parent and child are both affected with SLE and are the only affected family members participating; Simplex: one participant affected with SLE and at least one first-degree unaffected family member participating; Sporadic case: only one participant affected with SLE and no other family members participating. The bars show the number of pedigrees and affected individuals with SLE in each category that are enrolled and useful for genotyping.

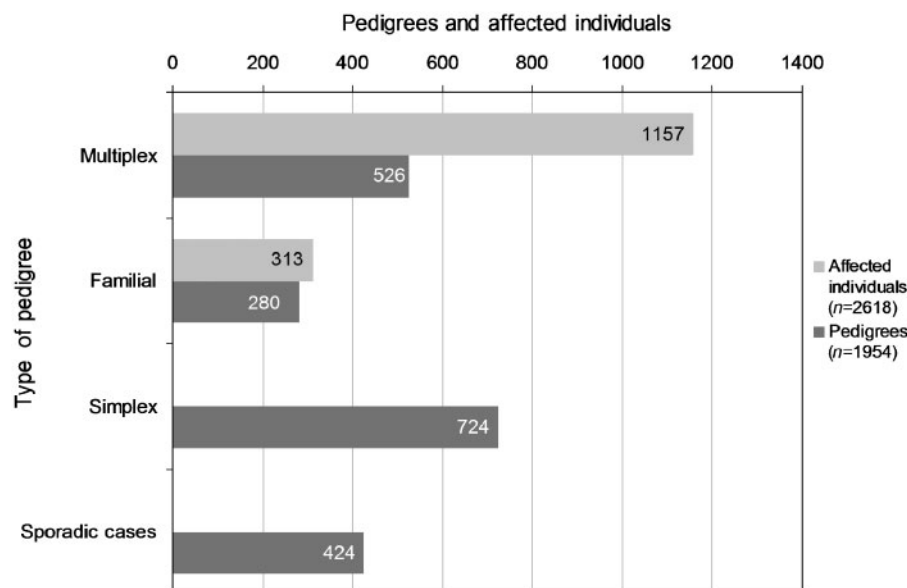


Fig. 2 Race and ethnicity of families enrolled in LFRR. **(A)** Number of individuals affected with SLE by race and ethnicity. **(B)** Number of pedigrees enrolled by race and ethnicity. The racial and ethnicity classification is based on the National Institutes of Health (NIH) guidelines for studies recruiting minorities (Office of Management and Budget, 2001; NOT-OD-01-053). These include two ethnic categories: Hispanic or Latino, and Not Hispanic or Latino, and five racial categories: American Indian or Alaska Native, Asian, Black or African-American, Native Hawaiian or Other Pacific Islander and White. The Gullah are a group of African-American families from the Sea Islands of South Carolina that have remained relatively isolated since their arrival from West Africa in the early 1700s and are of great interest in lupus genetic studies given their genetic homogeneity in comparison to other African-American communities.

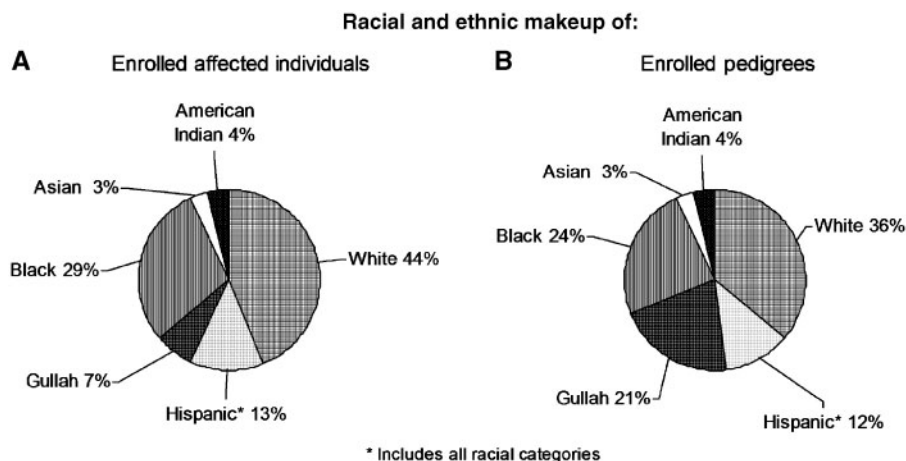
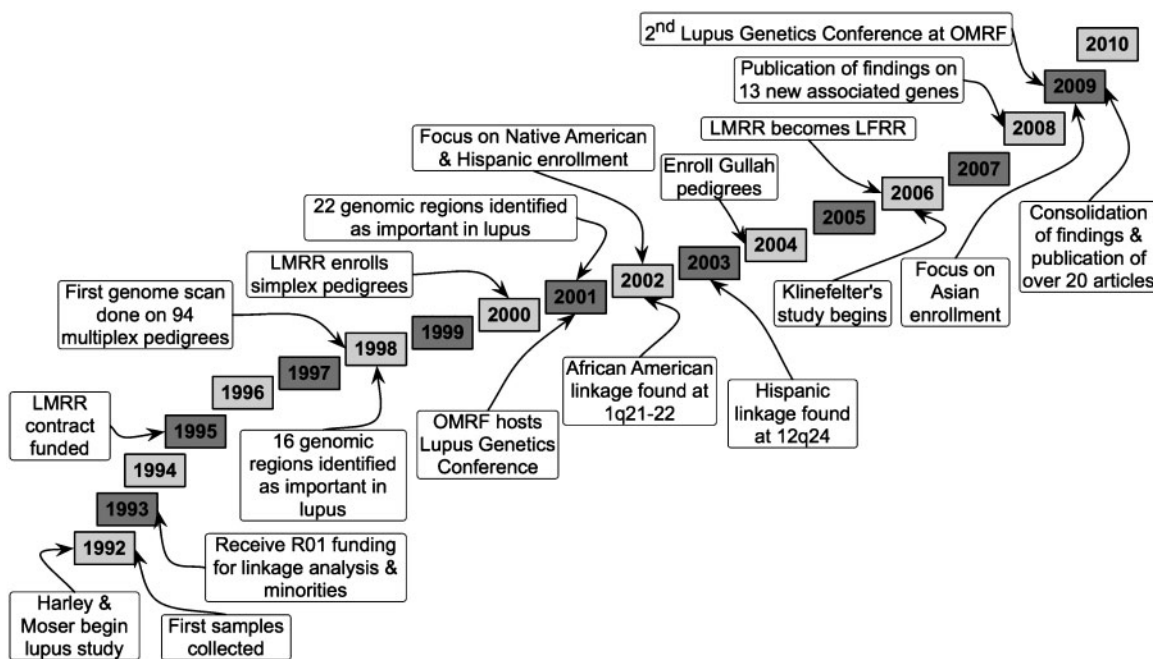


Fig. 3 Timeline of LFRR. Work on assembling and distributing materials and data from families with one or more living members diagnosed with lupus started in 1991 and has led to a large number of discoveries and scientific publications, as well as the world's largest repository of SLE families in 2010.



Potential participants placed into pending (whose eligibility is determined by reviewing medical records before obtaining a blood sample) are provided with an explanation of why they are not proceeding immediately to enrolment, as well as release of information (ROI) forms

and the OMRF Notice of Privacy Practices. Recruiters contact pending participants monthly regarding the status of records acquisition and review. After record review, enrolment or removal procedures follow based on the decision of the reviewer.

All enrolled participants who complete participation receive a small reimbursement, a copy of their serology report (if desired) and a free subscription to the annual LFRR newsletter, the *Lupus Linkage Newsletter*.

Medical records acquisitions

The clinical data that are collected in the database are based on verifiable medical records. The records pertinent to the diagnosis and treatment of each affected participant are obtained by records acquisition specialists after the participant has sent a signed Health Insurance Portability and Accountability Act (HIPAA)-compliant release of information form along with a list of facilities that will provide the records. Each health care provider is contacted in writing, specifying the patient and time frame for records of interest and a description of the purpose of the request, and provided a copy of the signed ROI form. Collecting all the records often requires repeated contacts with the providers; the database helps track the requests and monitors ROI expiration dates. Minimizing the expenses of medical record retrieval is a continuing goal.

Medical record abstraction

The medical record reviewer (M.D., P.A.-C. or R.N. specifically trained in this area) analyses the interview from the recruiter to determine whether the prospective participant is likely to meet ACR classification criteria for SLE [11, 12]. The medical records are screened by the reviewer and specific information is collected regarding which of the ACR classification criteria have been met and the earliest date that they were observed, as well as more than 200 additional laboratory and clinical data points that characterize the SLE manifestations and treatment of each affected participant.

Bio-specimen processing and storing

The biological samples of enrolled participants are usually received by overnight courier and immediately assigned a unique lupus genetics study barcode (LGScore) for sample and data-tracking purposes. The most commonly received bio-specimen is whole peripheral blood, however, occasionally mouthwash samples or preserved tissues are the source of DNA. Up to 66.5 ml of blood are collected from female participants; one additional 9.5 ml blood tube is collected from affected male participants (for karyotyping). An aliquot of blood is sent to the Clinical Laboratory Improvement Amendments (CLIA)-approved Clinical Immunology Laboratory at OMRF, where a standardized set of serological tests are performed. These include an ANA test on HEP-2 cells (an indirect fluorescent antibody test); anti-dsDNA (with titre by immunofluorescence against *Crithidia luciliae*); extractable antibodies (ENA) (by precipitin) with detection of antibodies against Ro/SSA, La/SSB, Smith (Sm), nRNP, ribosomal P, PM-Scl, Jo-1, Mi-2, Scl-70 and unidentified precipitins; aPLs (by ELISA) comprising IgG, IgM and IgA cardiolipin antibodies; and in the case of affected participants, an additional aliquot of 100 µl is used for determination of the total haemolytic complement (CH-50).

The main components isolated from blood for the bio-specimen bank include serum, plasma, peripheral blood mononuclear cells (PBMCs) for EBV-transformed cell lines (EBVTLs) and frozen specimens, and granulocytes for DNA isolation. Each of the components is stored in colour-coded aliquots of different volumes that allow for both efficient sample retrieval and minimization of freeze/thaw cycles. Dilutions of DNA are stored in 96-well deep-well storage plates at different dilutions to facilitate distribution to approved users (detailed handling protocols are available as supplementary data at *Rheumatology Online*).

The physical specimens collected from each participant are divided into two sets and are housed in two independent buildings that are both physically separated by space and electrical control, and monitored by onsite and offsite systems. This arrangement allows for the availability of backup samples for each participant in case of a catastrophic failure at one location.

Data recording and quality control

The data collected from each participant include the interview, questionnaire, serology results from the OMRF Clinical Immunology Laboratory and medical record review. These data are entered into the database; the questionnaires are scanned and the serology is directly imported from the computer system of the OMRF Clinical Immunology Laboratory. The remaining information is manually entered by a specialized member of the LFRR staff and later checked for accuracy by a second member.

Information technology

The informatics infrastructure is critical to operate, maintain and store data for a successful registry. Day-to-day issues including database schemas, storage requirements and informatics integration with enrolment workflows are indispensable. The two central canons of the informatics approach to the challenges posed by the LFRR are that (i) the accurate identification of subjects and samples is paramount and (ii) we must automate every possible procedure and process.

Every participant and enrollee is related to a central database via a single barcode identification scheme, the LGScore, which is a six-digit identifier. The database server is a relational database, wherein different tables of data are interconnected or related to one another. The MySQL relational database management system has more than 200 tables and 1700 fields, which are available across the network to approved users. The main LFRR interface for staff contains a number of custom forms and code currently housed in a Microsoft Access front end.

The LFRR inventory allows the accurate and precise location of the biological samples geographically in freezers by modelling them within the virtual world of the computer database. Each aliquot is linked back by its LGScore to its respective study subject and sample generation event, the precise coordinates of location

in storage and any addition, removal or manipulation. The complete inventory of the registry can be appraised and followed in real time, down to the volume of each aliquot.

The genotypic data collected by the LFRR are also processed through the informatics system, which has participated in the analysis of more than 650 000 SNP markers. More than 1 billion genotypes are stored and distributed by a Web-based interface for easy export.

Finally, the informatics team is also responsible for quality control and security of the data. The latest technology is applied to encryption systems, anti-virus screening, firewalls, secured data centre, software patches, multiple checks on data entry, software-level data validation and database-level data constraints.

The information technology (IT) team provides custom-made solutions for the needs of the LFRR staff. By utilizing open standards, mostly free or open source software and avoiding proprietary systems that hold the danger of locking the LFRR into a single vendor or system, we are able to grow and adapt as needed without unnecessary purchasing expenses (detailed information available as supplementary data, available at *Rheumatology* Online).

Institutional review board

Participant consent and privacy are the top priorities of the LFRR. The LFRR operates under the guidance of two institutional review boards (IRBs)—OMRF and the University of Oklahoma Health Sciences Center (OUHSC). The scientific activity requires a full-time IRB coordinator within LFRR, who responds to human subject and privacy-related questions and concerns from the LFRR recruiters, referral sites, users and collaborators. The coordinator oversees all communications between the LFRR and the IRBs, enforces compliance with human subject rules and regulations and supervises the successful completion of the Collaborative Institutional Training Initiative (CITI) Course for the Protection of Human Subjects in Research for all new employees.

All of the forms sent to participants, including the informed consent forms, the LFRR Web site text, advertising documents and revisions or modifications of any of these require IRB approval. An average of 45 new and revised proposals requiring IRB approval have been submitted every year in recent years. In addition, both IRBs require annual continuing review progress reports.

The addition of investigators at other institutions as LFRR affiliate sites also requires approval by the IRBs of the affiliate site and of OMRF and OUHSC; each site uses an informed consent document designed to satisfy the regulations of the IRBs of participating institutions.

Distribution of LFRR materials to external scientists

One of the main goals of the LFRR is to make the wealth of clinical and biological resources available to the relevant scientific community. It is important to note that the data and material of the LFRR are suitable for addressing some scientific problems but may not be a good resource for others; one of the major reasons for writing this article was to help interested scientists assess whether their scientific

questions of interest can be addressed with the data or materials in the LFRR (Fig. 4).

The application process to become an approved scientific user of the LFRR includes completing the application packet, agreeing to the letter of understanding by the investigator and responsible institutional official, and approval by the IRB of the requesting institution. This information is reviewed by the LFRR Scientific Advisory Committee (SAC), the OMRF Director of Research Administration and the NIAMS LFRR Program Officer. Approval is either granted or denied; often, additional information is required. (Application packet is available in supplementary materials at *Rheumatology* Online and at www.lupus.omrf.org/LFRRApp.html.)

The available data include a database release file, genotyping results from about 300 loci from pedigrees informative for linkage, and pedigree structure diagrams for all pedigrees except sporadic cases. The biological samples that can be requested include DNA, serum and plasma and a renewable source of DNA, which are the primary cultures of transformed B lymphocytes.

Discoveries derived from LFRR data

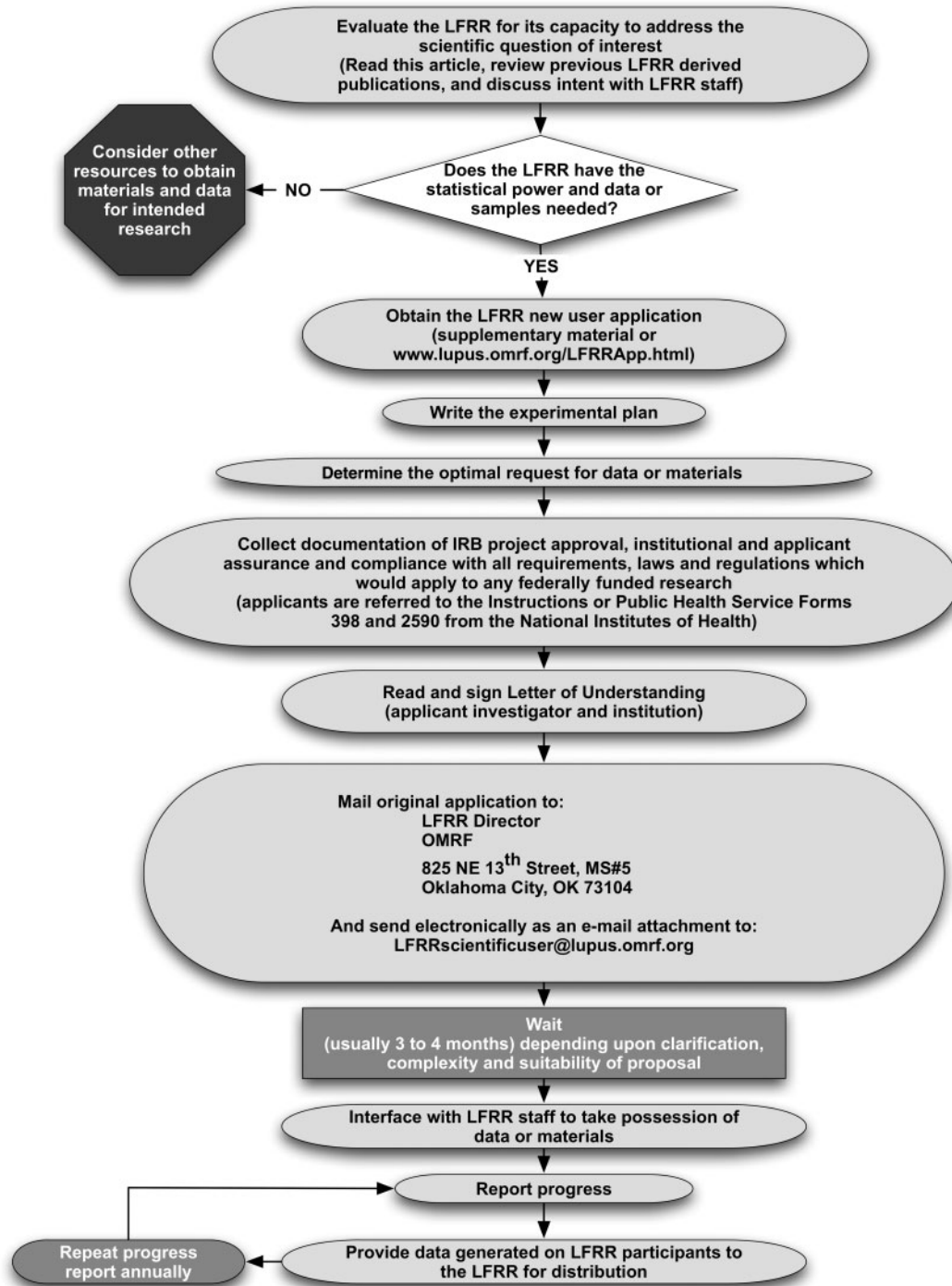
To date, more than 100 peer-reviewed scientific publications have resulted from the samples and data collected in the LFRR. We are optimistic that the body of work resulting from the LFRR will continue growing at an ever-increasing pace, paralleling the dramatic advance in genetic technologies.

Some of the early discoveries resulting from LFRR were derived from the first genome scan on 94 multiplex families completed in 1998. This study identified a genetic linkage to chromosome 1q21–22 near the *FcγRIIIA* locus, which was much stronger in African-American pedigrees than European-American pedigrees [13]. Meanwhile, linkage to 13q32 was identified in African-American families containing male SLE cases [14], to 4p16–15.2 in European-American but not African-American pedigrees [15], and to 12q25 in Hispanic and European pedigrees [16].

A meta-analysis of linkage studies revealed the most powerful effect on chromosome 6 in the region containing HLA and the second strongest effect on chromosome 16 [17]. The appreciation of the relative magnitude of this linkage fostered a candidate gene study that not only identified *ITGAM* (CD11b, CR3) as an associated gene, but also presented, using trans-ancestral mapping, R77H as the variant of *ITGAM* that is responsible for this association [18, 19].

In 2003, evidence was found that the risk of SLE in men with Klinefelter's syndrome (47, XXY) is similar to the risk in 46, XX women, and approximately 14 times higher than in 46, XY men, consistent with the notion that SLE predisposition is partly due to an X chromosome gene-dose effect [20], a finding replicated in a mouse model of SLE [21]. The groundbreaking International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEM) genome-wide association studies (GWAS) study that identified 13 new genetic effects in lupus [22] had the LFRR samples as its foundation; nearly all of these effects

Fig. 4 Step-by-step instructions for a serious scientist wanting LFRR data or material.



have subsequently been confirmed. Indeed, at this time there are 37 genes known to be associated with SLE (Table 1). Of these, the LFRR has contributed to the first discovery or to the all-important replication and confirmation of nearly 30 of these associations (in published and unpublished data).

Strengths and weaknesses of the LFRR

The LFRR is a unique resource that has adapted to changing opportunities (e.g. multiplex to simplex pedigrees, discussed above). In addition, many problems have been encountered and sometimes difficult choices have

TABLE 1 Published SLE gene associations significant at $P < 5 \times 10^{-8}$ as of March 2010

Gene	Published P-value	Published OR	Race	Reference	Confirmed reference
<i>PTPN22</i> ^a	$<9 \times 10^{-5}$ (EU)	1.4	EU	[36]	[22], many
<i>FcGR2A-FcGR3A</i> ^a	$<6.78 \times 10^{-7}$ (EU)	0.74	EU, AA, AS	[22]	
<i>FcGR3B</i>	2.7×10^{-8} (EU)	–	EU, AS	[37]	
<i>CRP</i> ^a	$<6.41 \times 10^{-7}$ (AA)	0.49	EU, AA	[59]	
<i>NMNAT2</i> ^a	1.08×10^{-7} (EU)	0.85	EU	[22]	[27]
<i>TNFSF4</i>	$<6.08 \times 10^{-7}$ (EU)	–	EU, AS, HI	[38]	[32, 38–41]
<i>IL10</i>	4.0×10^{-8} (EU)	1.19	EU	[27]	
<i>HIC2-UBE2L3</i> ^a	7.53×10^{-8} (EU)	1.22	EU, AS	[22]	[32]
<i>RASGRP3</i> [*]	1.3×10^{-15} (AS)	0.7	AS	[32]	
<i>STAT1, STAT4</i> ^a	1.9×10^{-9} (EU)	1.55	EU, AS, HI	[42]	[22, 28, 30, 32, 43–46]
<i>PXK</i> ^a	7.10×10^{-9} (EU)	1.25	EU	[22]	[27, 28]
<i>TREX1</i> [*]	$<4.1 \times 10^{-7}$ (EU)	25	EU	[65]	
<i>BANK1</i>	3.7×10^{-10} (EU)	1.4	EU, AA, AS, HI	[47]	[28, 39, 41]
<i>TNIP1</i>	3.8×10^{-13} (EU)	1.27	EU, AA, AS, HI	[32]	[27]
<i>UHRF1BP1</i>	2.22×10^{-8} (EU)	1.17	EU, AS	[27]	
HLA-DR2 and DR3 ^a	1.71×10^{-52} (EU)	2.36	EU	[22]	Many
Complement components C1q, C4A&B, C2		5–10	EU, AS		[60–64]
<i>PRDM1, ATG5</i> ^a	1.74×10^{-8} (EU)	1.19	EU	[22]	[27]
<i>TNFAIP3</i> ^a	2.9×10^{-12} (EU)	2.3	EU, AS	[48]	[32, 39, 49–52]
<i>IKZF1</i>	2.75×10^{-23} (AS)	0.72	AS	[32]	
<i>JAZF1</i>	1.5×10^{-9} (EU)	1.19	EU	[27]	
Multiple genes [*]	2.12×10^{-8} (AS)	0.83	AS	[32]	
<i>IRF5</i> ^a	4.4×10^{-16} (EU)	1.45	EU, AA, AS, HI	[53]	[22, 30, 32, 39, 44, 53–58], many
<i>XKR6</i> ^a	2.51×10^{-11} (EU)	1.23	EU, AS	[22]	
<i>BLK</i> ^a	7.0×10^{-10} (EU)	1.22	EU, AA, AS, HI	[22]	[28, 30, 32, 39, 45]
<i>LYN</i> ^a	5.4×10^{-9} (EU)	0.77	EU	[22]	[29]
<i>LRRC18-WDFY4</i>	7.22×10^{-12} (AS)	1.24	AS	[32]	[39]
<i>KIAA1542/PHRF1</i> ^a	3.0×10^{-10} (EU)	0.78	EU, AA	[22]	[27, 28]
Multiple genes [*]	1.25×10^{-16} (AS)	1.29	AS	[32]	
<i>ETS1</i>	1.77×10^{-25} (AS)	1.37	AS	[32]	[39]
<i>SLC15A4</i>	1.77×10^{-11} (AS)	1.26	AS	[32]	
<i>ITGAM-ITGAX</i> ^a	1.61×10^{-23} (EU)	1.62	EU, AS, HI	[22]	[18, 30, 31]
<i>IRAK1, MECP2</i> ^a	1.2×10^{-8} (EU, AS)	1.39	EU, AS, HI	[33]	[34, 35]

^aAssociation in LFRR samples has been published; ^{*}association in LFRR samples does not reach $P < 0.01$; some probabilities (<) are known to satisfy genome-wide significance from other publications or unpublished data. EU: European; AA: Black; AS: Asian; HI: Hispanic; OR: odds ratio.

been made in hopes of optimally applying resources. For example, the LFRR uses its resources to expand the collection instead of attempting periodic reassessment of previously enrolled participants. This means that the collection is larger at the cost of having very little data on the longitudinal disease course of SLE cases or on family members that may have subsequently developed SLE (or another autoimmune disease) (Tables 2 and 3).

A major logistic challenge encountered in the establishment of registries is the accurate collection of biometric registry entry information [23]. A unique feature of the data collection in the LFRR is that the clinical features of the affected individuals are available not only in qualitative form, but we also use a scoring system that allows any user of the data to evaluate the level of certainty with which each data point was obtained [10].

For example, individuals who have apparently convincing SLE based on their interviews and questionnaires, but for whom the available medical records are insufficient to document four or more ACR criteria, are labelled as affected with limited evidence. This allows individual scientists to determine how stringent they want to be in the inclusion or exclusion of a putative case for their study. We also record the level of evidence confirmed by medical records for each individual ACR criterion, again enabling researchers to make their study group more or less homogeneous based on their particular needs. (The scoring system is available in supplementary materials available at *Rheumatology* Online and published previously [10, 11, 12].)

Information about race, ethnicity or ancestry is collected using the revised guidelines established by the Office of

TABLE 2 Strengths of the LFRR

Strengths
<p>Large collection including:</p> <ul style="list-style-type: none"> Multiplex families and trios (account for 29% enrolled pedigrees) Large proportion of minorities (56% of all participants) Extensive data collection on SLE affected individuals (more than 5 000 000 points of clinical, serologic, demographic and genotyping data) Scoring system for SLE ACR criteria using medical records and laboratory tests Customized database with tracking system of all biological and clinical data Standardized serology reports on 99% of samples <p>Large control cohort:</p> <ul style="list-style-type: none"> At least one control is recruited for each affected participant Data collection for controls is similar to data of SLE affected: questionnaires, serologic, demographic and genotyping data are recorded in similar manner <p>Quality control</p> <ul style="list-style-type: none"> Genotyping for race, relatedness and gender in addition to self-reports <p>Scoring system of ACR criteria allows for:</p> <ul style="list-style-type: none"> Assurance that each SLE participant meets at least four SLE ACR criteria Scientific users able to choose participants based on the number of criteria met (LFRR offers the same data for SLE cases with limited evidence, participants with more than four criteria confirmed by source documents) <p>Long-standing study that has overcome many obstacles while implementing the lessons learned from them (15 years worth)</p> <p>Storage of samples and data in two separate and independent physical locations with independent power supplies to provide added security</p> <p>Receives valuable support from:</p> <ul style="list-style-type: none"> National Institute of Arthritis, Musculoskeletal and Skin Diseases (NIAMS) Oklahoma Medical Research Foundation (OMRF) Other funding agencies and institutes [e.g. National Institute of Allergy and Infectious Diseases (NIAID), Alliance for Lupus Research (ALR)] Network of affiliate sites Lupus support groups and awareness groups Physicians across the USA as well as internationally <p>Patient questionnaires and data collection are amended for changing interests and requirements</p>

Management and Budget (OMB) at the NIH (2001) for studies recruiting minorities. These include two ethnic categories: Hispanic or Latino, and Not Hispanic or Latino; and five racial categories: American Indian or Alaska Native, Asian, Black or African-American, Native Hawaiian or Other Pacific Islander and White [NOT-OD-01-053]. Our strategy has been to enrol the participants based on self-referred race and ethnicity, but we complement this information with questions about their ancestors' countries of origin for the previous two generations. Relatedness within the family can sometimes be imprecise, so a pedigree diagram is drawn for each proband and cross-referenced to the information obtained from other participating family members and is then confirmed through genotyping.

Our genotyping studies have included a set of ancestry informative markers (AIMs) that are subject to principal component analysis (PCA) to identify misidentified ancestry, admixture and individuals that have a unique combination of ethnic backgrounds that make them outliers for population-specific studies. The genotyping also permits corroboration of relatedness within families and correct gender assignment. This level of quality control has resulted in a loss of ~6% of participants after genotyping, but these quality control steps ensure that the best possible subjects and subject material are being provided to their experiments.

Recruitment of ethnic minorities with SLE is a priority of the LFRR, with a continuing goal of enrolling 40% of participants of non-European ancestry. In early 2010, the collection was composed of 56% of ethnic minority affected (Fig. 2). The other two focus subgroups have been multiplex pedigrees (27%) and affected males (11%) (Fig. 1 and supplementary figure available at *Rheumatology* Online). The availability of these more difficult to obtain subjects makes LFRR a unique resource.

Future direction of LFRR

After almost three decades, lupus genetics is nearing mid-life, at least for the discovery of genetic effects. In 2010, 37 genes are known to contribute to SLE, mainly in European and Asian ancestries. There must be another 50 genes that will become established with the exploration of other samples or further fine-mapping. Since these considerations do not include the genes that will be found after a careful exploration of other ancestries (especially, African and American-Indian), we predict that well over 100 genes will be convincingly associated with SLE in another decade. Each one of these genes will contribute to the mechanisms causing lupus in its own particular way.

Additional genes, especially those with minor-risk alleles (allele frequencies ≤ 0.05) and recessive inheritance

TABLE 3 Weaknesses and solutions of the LFRR

Weaknesses	Solutions
Not longitudinal No disease activity or damage indices	Patient information available for potential follow-up studies Clinical information available in database to calculate surrogate measures
Asians are under-represented	More resources are being dedicated to recruiting Asian families and controls
Changing concept of ethnicity and race: Hispanic was once considered a race, but is now an ethnicity, therefore Hispanics recruited before 2001 often do not have race recorded	The race and ethnicity are self-reported by the participants but verified through genotyping
Incomplete standardization of measurements and dates make international participation difficult Lack of face-to-face recruitment	Errors have been rectified in the standard operating procedures for international measurements and dates LFRR recruiters attend several health- or lupus-related events each year in an effort to increase face-to-face interactions with participants

represent an untapped source of genetic information that remains to be explored with the new genotyping and next-generation DNA-sequencing technologies. Advances in this field may prove that the estimate of only 100 genes is conservative.

The recurrent themes in the genes discovered will direct biologic work on mechanisms, holding the promise of important critical progress through basic understanding. Two such examples are the role of complement in age-related macular degeneration (366 publications since the publication of the GWAS-based discovery of association with complement Factor H in 2005 [24]), and autophagy in inflammatory bowel disease (83 papers published since the appreciation by GWAS of autophagy's role in the pathogenesis of Crohn's disease [25]).

The scientific goals that we envision for the LFRR are clear and diverse. We are embarking on fine-mapping of 35 of the 37 established genes associated with SLE. At the same time, we are performing GWASs in cases and controls of American-Indian, Asian and African ancestries with the goal of dissecting ancestry-specific effects and of using ancestry to help identify the polymorphisms responsible for disease risk.

Next-generation sequencing is anticipated to identify rare variants associated with lupus, which would be missed by the GWAS common-disease common-variant model. We expect that there will be a return to family studies for the identification of recessive risk variants. There are also experiments under way to understand the role of epigenetic mechanisms such as DNA methylation and microRNA dysregulation in SLE pathogenesis, and we will also make specific excursions into the world of genomic rearrangements, such as copy number variants (CNVs), inversions and indels, in an effort to be comprehensive when trying to find causal variants.

What started as a small lupus study with a staff of three has burgeoned into a multi-institutional worldwide collaborative team of scientists dedicated to the shared goal of understanding the genetic causes of SLE for the purpose of discovering better diagnostics and developing safer and more effective treatments. In 2010, the LFRR remains

the largest collection of lupus multiplex and simplex pedigrees in the world and continues to offer its biological and clinical data to scientists whose research will further contribute to unraveling the lupus puzzle.

Rheumatology key messages

- The LFRR is an important resource for understanding the genetics of lupus.
- Scientist can address carefully formulated scientific questions by requesting LFRR materials and data.
- The LFRR will be a useful contributor to lupus genetics for many years to come.

Acknowledgements

The authors thank additional contributors to the LFRR effort: Sharon Johnson, Summer Frank, Amy Butler, Barbara Leatherwood, Patti Grounds, Jacey Bush, Dominique Williams, Adriana Rojas-Villarraga, Lina Amezcuita, Ryan Parker, Brandon Scharrer, Tamiko Cabatic, Lauren Evans, Michelle Calvo, Nicole Weber, Kay Davis, Sarah Dawson, Kurt Downing and Neeraj Asundi. We thank the patients and all the other referring health care providers.

Funding: This work was supported by the National Institutes of Health, mainly from the National Institute of Arthritis, Musculoskeletal and Skin Diseases (N01AR62277) with important contributions from additional grants from the National Institutes of Health (R37AI24717, R01AR42460, P20RR020743, R01AR053734, P01AR049084, P20AR046669, RC1AR058554 and R01AR043274). G.S.G. and D.L.K. were supported by the South Carolina Clinical & Translational Research Institute, Medical University of South Carolina's Clinical and Translational Science Award (CTSA) and National Institutes of Health/National Center for Research Resources (UL1RR029882). L.R.E. was funded by the Board of Regents of the State of Louisiana.

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

- Lawrence JS, Martins CL, Drake GL. A family survey of lupus erythematosus. 1. Heritability. *J Rheumatol* 1987;14: 913–21.
- Alarcón-Segovia D, Alarcón-Riquelme ME, Cardiel MH *et al*. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. *Arthritis Rheum* 2005;52:1138–47.
- Deapen D, Escalante A, Weinrib L *et al*. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis Rheum* 1992;35:311–8.
- Block SR, Winfield JB, Lockshin MD, D'Angelo WA, Christian CL. Studies of twins with systemic lupus erythematosus. A review of the literature and presentation of 12 additional sets. *Am J Med* 1975;59:533–52.
- Reichlin M, Harley JB, Lockshin MD. Serologic studies of monozygotic twins with systemic lupus erythematosus. *Arthritis Rheum* 1992;35:457–64.
- Reva CL, Charles GH, Frank CA *et al*. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* 1998;41: 778–99.
- Seldin MF, Qi L, Scherbarth HR *et al*. Amerindian ancestry in Argentina is associated with increased risk for systemic lupus erythematosus. *Genes Immun* 2008;9:389–93.
- Hopkinson ND, Doherty M, Powell RJ. Clinical features and race-specific incidence/prevalence rates of systemic lupus erythematosus in a geographically complete cohort of patients. *Ann Rheum Dis* 1994;53:675–80.
- Seligman VA, Lum RF, Olson JL, Li H, Criswell LA. Demographic differences in the development of lupus nephritis: a retrospective analysis. *Am J Med* 2002;112: 726–9.
- Rao SA, Olson JM, Moser KL *et al*. Linkage analysis of human systemic lupus erythematosus-related traits: a principal component approach. *Arthritis Rheum* 2001;44: 2807–18.
- Tan EM, Cohen AS, Fries JF *et al*. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982;25:1271–7.
- Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40: 1725.
- Edberg JC, Langefeld CD, Wu J *et al*. Genetic linkage and association of Fcγ receptor IIIA (CD16A) on chromosome 1q23 with human systemic lupus erythematosus. *Arthritis Rheum* 2002;46:2132–40.
- Xing C, Gray-McGuire C, Kelly JA *et al*. Genetic linkage of systemic lupus erythematosus to 13q32 in African American families with affected male members. *Hum Genet* 2005;118:309–21.
- Gray-McGuire C, Moser KL, Gaffney PM *et al*. Genome scan of human systemic lupus erythematosus by regression modeling: evidence of linkage and epistasis at 4p16-15.2. *Am J Hum Genet* 2000;67:1460–9.
- Nath SK, Quintero-del-Rio AI, Kilpatrick J, Feo L, Ballesteros M, Harley JB. Linkage at 12q24 with systemic lupus erythematosus (SLE) is established and confirmed in Hispanic and European American families. *Am J Hum Genet* 2004;74:73–82.
- Lee YH, Nath SK. Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. *Hum Gene* 2005;118:434–43.
- Nath SK, Han S, Kim-Howard X *et al*. A nonsynonymous functional variant in integrin-α-M (encoded by ITGAM) is associated with systemic lupus erythematosus. *Nat Genet* 2008;40:152–4.
- Han S, Kim-Howard X, Deshmukh H *et al*. Evaluation of imputation-based association in and around the integrin-α-M (*ITGAM*) gene and replication of robust association between a non-synonymous functional variant within *ITGAM* and systemic lupus erythematosus (SLE). *Hum Mol Genet* 2009;18:1171–80.
- Scotfield RH, Bruner GR, Namjou B *et al*. Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. *Arthritis Rheum* 2008;58: 2511–7.
- Smith DL, Dong X, Du S, Oh M, Singh RR, Voskuhl RR. A female preponderance for chemically induced lupus in SJL/J mice. *Clin Immunol* 2007;122:101–7.
- Harley JB, Alarcón-Riquelme ME, Criswell LA *et al*. International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in *ITGAM*, *PXK*, *KIAA1542* and other loci. *Nat Genet* 2008;40:204–10.
- Brenner H, Schmidtman I, Stegmaier C. Effects of record linkage errors on registry-based follow-up studies. *Stat Med* 1997;16:2633–43.
- Klein RJ, Zeiss C, Chew EY *et al*. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005;308:385–9.
- Parkes M, Barrett JC, Prescott NJ *et al*. Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;39:830–2.
- Harley JB. IL-7R-α and multiple sclerosis risk. *Nat Genet* 2007;39:1053–4.
- Gateva V, Sandling JK, Hom G *et al*. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1228–33.
- Suarez-Gestal M, Calaza M, Endreffy E *et al*. Replication of recently identified systemic lupus erythematosus genetic associations: a case-control study. *Arthritis Res Ther* 2009;11:R69.
- Lu R, Vidal GS, Kelly JA *et al*. Genetic associations of LYN with systemic lupus erythematosus. *Genes Immun* 2009; 10:397–403.
- Hom G, Graham RR, Modrek B *et al*. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N Engl J Med* 2008;358:900–9.

- 31 Yang W, Zhao M, Hirankarn N *et al.* ITGAM is associated with disease susceptibility and renal nephritis of systemic lupus erythematosus in Hong Kong Chinese and Thai. *Hum Mol Genet* 2009;18:2063–70.
- 32 Han JW, Zheng HF, Cui Y *et al.* Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1234–7.
- 33 Sawalha AH, Webb R, Han S *et al.* Common variants within MECP2 confer risk of systemic lupus erythematosus. *PLoS One* 2008;3:e1727.
- 34 Jacob CO, Zhu J, Armstrong DL *et al.* Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus. *Proc Natl Acad Sci USA* 2009;106:6256–61.
- 35 Jacob CO, Reiff A, Armstrong DL *et al.* Identification of novel susceptibility genes in childhood-onset systemic lupus erythematosus using a uniquely designed candidate gene pathway platform. *Arthritis Rheum* 2007;56:4164–73.
- 36 Kyogoku C, Langefeld CD, Ortmann WA *et al.* Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am J Hum Genet* 2004;75:504–7.
- 37 Fanciulli M, Norsworthy PJ, Petretto E *et al.* FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat Genet* 2007;39:721–3.
- 38 Cunninghame Graham DS, Graham RR, Manku H *et al.* Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. *Nat Genet* 2008;40:83–9.
- 39 Yang W, Shen N, Ye DQ *et al.* Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS Genet* 2010;6:e1000841.
- 40 Delgado-Vega AM, Abelson AK, Sánchez E *et al.* Replication of the TNFSF4 (OX40L) promoter region association with systemic lupus erythematosus. *Genes Immun* 2009;10:248–53.
- 41 Chang YK, Yang W, Zhao M *et al.* Association of BANK1 and TNFSF4 with systemic lupus erythematosus in Hong Kong Chinese. *Genes Immun* 2009;10:414–20.
- 42 Remmers EF, Plenge RM, Lee AT *et al.* STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 2007;357:977–86.
- 43 Taylor KE, Remmers EF, Lee AT *et al.* Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. *PLoS Genet* 2008;4:e1000084.
- 44 Abelson AK, Delgado-Vega AM, Kozyrev SV *et al.* STAT4 associates with systemic lupus erythematosus through two independent effects that correlate with gene expression and act additively with IRF5 to increase risk. *Ann Rheum Dis* 2009;68:1746–53.
- 45 Yang W, Ng P, Zhao M *et al.* Population differences in SLE susceptibility genes: STAT4 and BLK, but not PTK, are associated with systemic lupus erythematosus in Hong Kong Chinese. *Genes Immun* 2009;10:219–26.
- 46 Namjou B, Sestak AL, Armstrong DL *et al.* High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups. *Arthritis Rheum* 2009;60:1085–95.
- 47 Kozyrev SV, Abelson AK, Wojcik J *et al.* Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nat Genet* 2008;40:211–6.
- 48 Graham RR, Cotsapas C, Davies L *et al.* Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nat Genet* 2008;40:1059–61.
- 49 Musone SL, Taylor KE, Lu TT *et al.* Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. *Nat Genet* 2008;40:1062–4.
- 50 Bates JS, Lessard CJ, Leon JM *et al.* Meta-analysis and imputation identifies a 109 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations. *Genes Immun* 2009;10:470–7.
- 51 Cai LQ, Wang ZX, Lu WS *et al.* A single-nucleotide polymorphism of the TNFAIP3 gene is associated with systemic lupus erythematosus in Chinese Han population. *Mol Biol Rep* 2010;37:389–94.
- 52 Shimane K, Kochi Y, Horita T *et al.* The association of a nonsynonymous single-nucleotide polymorphism in TNFAIP3 with systemic lupus erythematosus and rheumatoid arthritis in the Japanese population. *Arthritis Rheum* 2010;62:574–9.
- 53 Graham RR, Kozyrev SV, Baechler EC *et al.* A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* 2006;38:550–5.
- 54 Cunninghame Graham DS, Manku H, Wagner S *et al.* Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation. *Hum Mol Genet* 2007;16:579–91.
- 55 Kozyrev SV, Lewén S, Reddy PM *et al.* Structural insertion/deletion variation in IRF5 is associated with a risk haplotype and defines the precise IRF5 isoforms expressed in systemic lupus erythematosus. *Arthritis Rheum* 2007;56:1234–41.
- 56 Graham RR, Kyogoku C, Sigurdsson S *et al.* Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci USA* 2007;104:6758–63.
- 57 Sigurdsson S, Göring HH, Kristjansdottir G *et al.* Comprehensive evaluation of the genetic variants of interferon regulatory factor 5 (IRF5) reveals a novel 5 bp length polymorphism as strong risk factor for systemic lupus erythematosus. *Hum Mol Genet* 2008;17:872–81.
- 58 Kelly JA, Kelley JM, Kaufman KM *et al.* Interferon regulatory factor-5 is genetically associated with systemic lupus erythematosus in African Americans. *Genes Immun* 2008;9:187–94.
- 59 Edberg JC, Wu J, Langefeld CD *et al.* Genetic variation in the CRP promoter: association with systemic lupus erythematosus. *Hum Mol Genet* 2008;17:1147–55.
- 60 Botto M, Walport MJ. C1q, autoimmunity and apoptosis. *Immunobiology* 2002;205:395–406.
- 61 Fielder AH, Walport MJ, Batchelor JR *et al.* Family study of the major histocompatibility complex in patients with systemic lupus erythematosus: importance of null alleles of C4A and C4B in determining disease susceptibility. *Br Med J (Clin Res Ed)* 1983;286:425–8.
- 62 Yang Y, Chung EK, Wu YL *et al.* Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus

- (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. *Am J Hum Genet* 2007;80:1037–54.
- 63 Sullivan KE, Petri MA, Schmeckpeper BJ, McLean RH, Winkelstein JA. Prevalence of a mutation causing C2 deficiency in systemic lupus erythematosus. *J Rheumatol* 1994;21:1128–33.
- 64 Namjou B, Gray-McGuire C, Sestak AL *et al.* Evaluation of C1q genomic region in minority racial groups of lupus. *Genes Immun* 2009;10:517–524.
- 65 Lee-Kirsch MA, Gong M, Chowdhury D *et al.* Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. *Nat Genet* 2007;39:1065–7.