Meta-analysis of genome-wide linkage scans for renal function traits

Madhumathi Rao1, Amy K. Mottl2, Shelley A. Cole3, Jason G. Umans4, Barry I. Freedman5, Donald W. Bowden5, Carl D. Langefield5, Caroline S. Fox6, Qiong Yang7, Adrienne Cupples7, Sudha K. Iyengar8, Steven C. Hunt9 and Thomas A. Trikalinos10

1Division of Nephrology, Tufts Medical Center, Boston, MA, USA, 2UNC Kidney Center, School of Medicine, University of North Carolina, Chapel Hill, NC, USA, 3Southwest Foundation for Biomedical Research, San Antonio, TX, USA, 4Department of Medicine, Obstetrics and Gynecology, General Clinical Research Center, Georgetown University Medical Center, Washington, DC, USA, 5Nephrology, Center for Human Genomics, and Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA, 6NHBLI’s Framingham Heart Study and Center for Population Science, National Heart, Lung, and Blood Institute, Framingham, MA, USA, 7Department of Biostatistics, Boston University, Boston MA, USA, 8Division of Genetic and Molecular Epidemiology, Case Western Reserve University, Cleveland OH, USA, 9Cardiovascular Genetics Division, Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA and 10Center for Clinical Evidence Synthesis, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, 02111 MA, USA

Correspondence and offprint requests to: Thomas A. Trikalinos; E-mail: ttrikalinos@tuftsmedicalcenter.org

Abstract

Background. Several genome scans have explored the linkage of chronic kidney disease phenotypes to chromosomal regions with disparate results. Genome scan meta-analysis (GSMA) is a quantitative method to synthesize linkage results from independent studies and assess their concordance.

Methods. We searched PubMed to identify genome linkage analyses of renal function traits in humans, such as estimated glomerular filtration rate (GFR), albuminuria, serum creatinine concentration and creatinine clearance. We contacted authors for numerical data and extracted information from individual studies. We applied the GSMA nonparametric approach to combine results across 14 linkage studies for GFR, 11 linkage studies for albumin creatinine ratio, 11 linkage studies for serum creatinine and 4 linkage studies for creatinine clearance.

Results. No chromosomal region reached genome-wide statistical significance in the main analysis which included all scans under each phenotype; however, regions on Chromosomes 7, 10 and 16 reached suggestive significance for linkage to two or more phenotypes. Subgroup analyses by disease status or ethnicity did not yield additional information.

Conclusions. While heterogeneity across populations, methodologies and study designs likely explain this lack of agreement, it is possible that linkage scan methodologies lack the resolution for investigating complex traits. Combining family-based linkage studies with genome-wide association studies may be a powerful approach to detect private mutations contributing to complex renal phenotypes.

Keywords: albuminuria; chronic kidney disease; glomerular filtration rate; linkage scans; meta-analysis

Introduction

Genetic factors contribute to the development and progression of chronic kidney disease (CKD). Familial aggregation of kidney disease has been shown in multiple epidemiologic studies [1–3], and traits such as estimated glomerular filtration rate (eGFR) and albuminuria are modestly heritable [4–6]. While the advent of large genome-wide association studies has increased the scope for the identification of kidney disease susceptibility genes [7], with the exception of mapping by admixture linkage disequilibrium (LD) [8–10], both association and linkage approaches have had limited success to date.

A genome-wide linkage scan is an unbiased approach to interrogate the genome with multiple markers that identify evenly spaced locations along each of the human chromosomes in families with multiple subjects manifesting the trait of interest. Linkage analysis tests for co-segregation of the marker with a trait or phenotype of interest within a family implying physical linkage between the marker and a phenotype locus. Since the first genome-wide linkage scan for nephropathy genes in 1998 [11], there have been many studies addressing a range of kidney phenotypes in a variety of study populations, ranging from the general population to populations with underlying comorbidities such as diabetes or hypertension which could predispose to CKD. Multiple linkage peaks have been observed with little overlap both for a single phenotype and for related CKD phenotypes. Many peaks have not reached genome-wide significance, possibly due to linkage and population heterogeneity, small sample sizes and limited power or because they were simply false positives. Some of the
major challenges to the linkage approach as applied to complex genetic diseases include the need for large numbers of relative pairs to have sufficient power to detect loci that confer modest effects and a relatively homogeneous population with regard to ethnicity, comorbidity and distribution of the trait [12, 13].

One approach to compare genetic evidence across linkage studies is a meta-analysis, preferably with the primary data from each study. We therefore undertook a collaborative genome scan meta-analysis (GSMA) [14–16] of existing genome-wide linkage studies of quantitative CKD traits including albuminuria (albumin creatinine ratio, ACR), eGFR, serum creatinine concentration and creatinine clearance. These measures exhibit a continuous phenotypic spectrum and may offer more power than categorization based on arbitrarily defined thresholds. The GSMA is an exploratory data analysis method, which aims to highlight regions containing susceptibility loci, where further studies can be focused.

Materials and methods

**Literature search and eligible whole genome scans**

We searched PubMed for whole genome scans on four CKD traits, namely (estimated) glomerular filtration rate, ACR (urinary albumin), serum creatinine and creatinine clearance using combinations of the corresponding medical subject heading terms and terms such as ‘genome scan’ or ‘genome search’ and kidney disease, including the above-mentioned CKD traits (full search strategy provided as Supplement). Searches were limited to the English language and humans and to the period 01 January 1998 to 29 December 2010. We complemented searches by perusing reference lists of eligible papers and narrative reviews. Reference lists of relevant original papers and narrative reviews were studied and experts in the field were contacted. Investigators were invited to contribute necessary data for analyses.

We included whole genome scans on the aforementioned four outcomes, provided that they used >300 markers, irrespective of the statistical analyses employed. We excluded studies that mapped only specific chromosomes or chromosomal regions (all 22 autosomes had to be analyzed), genome-wide association studies, studies with <10 families and studies in people with Mendelian forms of inherited kidney disease.

**Definition of outcomes**

We included studies that estimated eGFR using the four-variable Modification of Diet in Renal Diseases equation [17] and calculated creatinine clearance with the Cockroft–Gault equation [18]. We accepted all methods of serum creatinine measurement irrespective of whether assay calibration was described or not. Studies that analyzed albuminuria as the phenotype expressed it as the albumin excretion rate, normalized to urinary creatinine (ACR).

**Data abstraction**

Bibliographic information (first author, journal and year of publication), area of recruitment, whether participants were diseased or healthy, number of probands, number of families or sibpairs (as applicable), reported outcomes, number of microsatellite markers, the name of the commercial marker set and the statistical methods and software used in each publication were recorded. The unit of the analysis was the population stratum, defined by ‘racial’ descent (the Europeanexcluding Hispanic, African-American, American Indian, Hispanic). We retained subgroups from different recruitment locations as separate strata if they were analyzed separately in the primary studies. Potential overlaps between studies were clarified in communication with study authors.

All identified research teams were invited to participate in this collaborative meta-analysis. Participating teams contributed sufficient statistics for each of their population strata. Whenever such data were not provided, we extracted them from publications, by digitizing graphs of linkage scores per pter (p terminal) distance in each chromosome using specialized software and corroborating with results reported in the text and tables. The reliability and validity of this approach has been previously demonstrated [19]. Whenever multiple analyses were available, we chose multipoint LOD scores over other analyses. We also selected results at the longest follow-up. We chose the more adjusted models over less adjusted ones, in an effort to select the most comparable analyses across studies. Most studies did not report unadjusted results.

**Genome scan meta-analysis**

We synthesized information across genome scans using methods described previously [15, 16, 20, 21]. Briefly, GSMA analyze chromosomal regions, rather than specific markers. It aligns all 22 autosomes in the genome and breaks them into 120 regions (bins) of ~30 cM each (a cM is a unit of distance based on recombination rates). These regions (bins) are the units of analysis. Larger chromosomes have more bins than smaller ones (e.g. Chromosome 1 has 10 bins, as it is ~300 cM long, while Chromosomes 21 and 22 have only two bins each). We name bins according to their position in the corresponding chromosome, e.g. Bin 7.2 would be the second bin on Chromosome 7. In each study, we represent each bin by the most significant P-value (or equivalent statistic, such as the highest LOD peak) that corresponds to it.

GSMA uses the relative ranks of the P-values (or other equivalent statistics) from the linkage analysis over all 120 bins in each study, with the highest rank being the most significant bin (highest peak) and the lowest being the least significant bins. If several bins are tied, they are assigned a common rank. In the GSMA, the bin that has the highest mean rank across all studies is the one where (on average) most high peaks are located. To judge the statistical significance of a bin’s mean rank, we perform a Monte Carlo permutation test [20].

For each bin, we also quantified the heterogeneity in the ranks across available population strata, i.e. a measure of the dispersion of the ranks in the individual studies around their mean. Intuitively, when this dispersion is very low, there is agreement among all studies on the rank of the particular bin. We quantify this dispersion with the Q statistic [15, 16], which is the sum of the squared differences of ranks from their mean. The distribution of the Q statistic is not known; we therefore tested whether the Q statistic implied significantly low heterogeneity based on Monte Carlo permutations.

**Subgroup and sensitivity analyses**

For each outcome, the main analysis combined all available population strata. We also performed subgroup analyses by ethnicity, baseline health status (general population, diabetes, hypertension or either of the latter two). The eGFR and creatinine clearance are transformation of serum creatinine; therefore, we also performed a sensitivity analysis where we treated all three outcomes together. To avoid duplication of information, we preferred serum creatinine over eGFR over creatinine clearance, where more than one were reported. Results were very similar with all permutations of the preferred order of outcomes and are thus not reported in further detail. We also performed analyses excluding digitized data and using only data for which primary numerical results were available. Sensitivity analyses were carried out using alternate data from studies with more than one analysis (e.g. using results from analyses with fewer adjustments instead of the most adjusted models), but as their results were very similar, they are not reported.

**Software and analytic details**

All Monte Carlo analyses were run for 100 000 permutations. In main analyses, we weighted bin ranks by the square root of the total number of subjects in the corresponding stratum (study). Because the optimal weighting for GSMA is unclear, we also performed unweighted analyses. Inferences were performed at the $\alpha = 0.05$ level. Thus, we considered a P-value ≤0.05 as suggestive of significant, and a P-value ≤0.00042 (using Bonferroni corrections for 120 comparisons) as genome-wide significant. Graph digitizing was performed with Engauge Digitizer (version 2.12, Mark Mitchell, 2002). All analyses and graphs were performed in Intercooled Stata 8.2 (Stata Corp. College Station, TX), using the ‘hegsma’ C plugin for Stata (developed by author T.A.T.). The Stata plugin has been verified against the HEGESMA software [16] and also allows for missing bin values.

**Results**

We included 12 studies corresponding to 22 distinct population (ethnic descent) strata. Table 1 summarizes their characteristics and additional descriptive data is shown in
Table 1. Characteristics of included studies

<table>
<thead>
<tr>
<th>Study (ref.)</th>
<th>GFR</th>
<th>Cr</th>
<th>CrCl</th>
<th>ACR</th>
<th>Population</th>
<th>Eur</th>
<th>Af</th>
<th>His</th>
<th>AI</th>
<th>Markers (set)</th>
<th>Analysis (software)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAFHSb [22, 23]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>General population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>848</td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>417 (Marshfield)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FHS [4, 5]</td>
<td>Y</td>
<td>Y</td>
<td>–</td>
<td>Y</td>
<td>General population</td>
<td>1055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>401 (Marshfield 9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1183;</td>
<td></td>
<td>1161;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1153</td>
<td></td>
</tr>
<tr>
<td>Utah [26]</td>
<td>–</td>
<td>Y</td>
<td>Y</td>
<td>–</td>
<td>General population</td>
<td>850</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NPL (GENEHUNTER)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eurosann [27]</td>
<td>–</td>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>General population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIND [28, 29]</td>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>Y</td>
<td>Diabetics</td>
<td>119</td>
<td></td>
<td>218</td>
<td>469</td>
<td>80</td>
<td>HE (SIBPAL)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joslinb [30]</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Y</td>
<td>Diabetics</td>
<td>857</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAFDGSb [31]</td>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>Y</td>
<td>Diabetics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AADMb [32]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>–</td>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>691</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHS [33]</td>
<td>Y</td>
<td>–</td>
<td>Y</td>
<td>–</td>
<td>Diabetics, hypertensives</td>
<td>902</td>
<td></td>
<td>165</td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperGEN [34]</td>
<td>Y</td>
<td>–</td>
<td>Y</td>
<td>–</td>
<td>Hypertensives</td>
<td>1040</td>
<td></td>
<td>1234</td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
<tr>
<td>GENOA [35]</td>
<td>Y</td>
<td>–</td>
<td>–</td>
<td>Y</td>
<td>Hypertensives</td>
<td>1022</td>
<td></td>
<td>1351</td>
<td></td>
<td></td>
<td>VC (SOLAR)</td>
</tr>
</tbody>
</table>

*As clarified in the Materials and methods section, different strata (by racial descent or within the same racial descent) were included as separate entries in the genome search meta-analysis. ‘–’ = not examined/not included, AADM, African-American Diabetes Mellitus Study; Af, African-American descent; AI, American Indian descent; Cr, serum creatinine; CrCl, creatinine clearance; CVD, cardiovascular disease; DHS, Diabetes Heart Study; DM, diabetes mellitus; GFR, estimated glomerular filtration rate; Eur, European descent; FIND, Family Investigation of Nephropathy and Diabetes; FHS, Framingham Heart Study; GENOA, Genetic Epidemiology Network of Arteriopathy; HE, Haseman-Elston regression; His, Hispanic (Mexican American) descent; HTN, hypertensive; HyperGEN, Hypertension Genetic Epidemiology Network Study; ND, no data; NPL, nonparametric linkage; SAFHS, San Antonio Family Heart Study; SAFDGS, San Antonio Family Diabetes/Gall Bladder Study; SHFS, Strong Heart Family Study; VC, variance components linkage analysis; STR, short tandem repeat markers; CIDR, Center for Inherited Disease Research.

*bMeta-analysis data extracted from digitized graphs.

*cTotal of 848 individuals in reference [23] or 486 individuals in reference [22].

*dTotal of 1224 individuals in reference [4] and 1055 individuals in reference [5].

*eData directly obtained from investigators.

*fThree populations from Oklahoma, Arizona and Dakota, respectively (USA) [25]. For GFR, the corresponding numbers are 1210, 1235 and 1220 [24].

*gPopulation enriched for CVD. There are three sequential longitudinal waves with 1516, 1193 and 850 analyzed subjects, respectively. Main analyses used the last wave, sensitivity analyses used the first or second wave instead.

*hThree isolated populations in Europe: MICROS study (Italy), ERF study (the Netherlands) and Vis study (Croatia).

*iFor map alignment, as the three studies had different genotypes, the deCODE genetic map was used as a reference to align individual maps so that a common linkage analysis could be performed on the same map, based on 1000 markers.

jThe CIDR genetic map is similar to the Marshfield Genetic map.

Supplementary Table 2. We obtained data directly from investigators for six studies (13 population strata) [4, 5, 24–26, 28, 29, 33, 34]. For the remaining studies, we extracted data from publications. There were 10 European-descent strata, 5 African-American (AA) strata, 3 Mexican American (MA) strata and 4 American Indian (AI) strata. Six studies had been carried out in general population cohorts and the remaining in populations enriched for diabetes or hypertension.

Estimated glomerular filtration rate

For the outcome of eGFR, we included data from 14 strata in our analysis. In weighted analyses that account for the size of each study, no bin reached genome-wide significance (P-value threshold ≤0.00042) (Figure 1, Table 2). Six bins (7.2, 7.4, 9.1, 13.4, 14.3 and 16.3, with corresponding chromosomal locations indicated in Table 2) reached suggestive significance (P-value threshold ≤0.05). There was no evidence for low between-scan heterogeneity (i.e. ranks were in agreement across population strata beyond what is expected by chance). Bins 9.1 and 7.2 had the smallest P-values for the mean rank (P = 0.0052 and P = 0.0062, respectively). Unweighted analyses were similar. Seven bins (the six from the weighted analyses and Bin 15.4) reached suggestive significance. Bins 16.3 and 9.1 had the lowest P-values (P = 0.0063 and P = 0.0102, respectively). Again, none of the bins with suggestive significance showed evidence for low between-scan heterogeneity.

Serum creatinine concentration

Weighted analyses among 11 genome scan strata resulted in 10 bins that reached suggestive significance (Figure 2, Table 2), although none reached genome-wide significance.
The lowest P-values for the mean rank were found for bins 9.6 and 7.3 (P = 0.00165 and 0.00285), respectively. Interestingly, five of six bins in the ninth chromosome (9.2 through 9.6) were among the 10 bins. None of the significant bins had evidence for low between-scan heterogeneity. Inferences in unweighted analyses were the same with the exception of Bin 11.4 which did not reach suggestive significance.

**Creatinine clearance**

For the outcome of creatinine clearance, four genome scan strata were included in the meta-analyses. In weighted analyses the same inferences were made with the exception of Bin 11.4 which did not reach suggestive significance.
analyses, five bins reached suggestive significance for mean rank (Figure 3; Table 2) with bins 7.3 and 6.6 having the lowest P-values ($P = 0.0195$ and 0.0310, respectively). No bin reached genome-wide significance. There was evidence for statistically significantly high between-scan heterogeneity with all three metrics for Bin 10.4. None of the significant bins had evidence for low between-scan heterogeneity. Inferences in unweighted analyses were the same with the exception of Bin 2.6 which did not reach suggestive significance.

**Albumin creatinine ratio**

The meta-analysis for ACR was based on 11 genome scan strata. Five bins reached suggestive significance (1.2, 3.6, 7.4, 7.5 and 10.4), but none achieved genome-wide levels and none had significantly low between-scan heterogeneity. The smallest P-values were found for bins 10.4 and 7.5 ($P = 0.0145$ and 0.0154, respectively; Figure 4, Table 2). Unweighted analyses were very similar, with six bins reaching suggestive significance (the aforementioned 5 and 20.2), again with no evidence for low between-scan heterogeneity.

**Subgroup analyses**

Figure 5 summarizes results of weighted main and subgroup analyses across all four phenotypes (the corresponding figure for unweighted analysis was similar—data not shown). Inferences on significance are not corrected for multiple comparisons. As shown in the figure, several bins were repeatedly identified as of suggestive statistical significance across subgroups and even across phenotypes. For example, Bin 7.4 reached suggestive significance in nine analyses for the outcomes of eGFR, ACR or serum creatinine concentration, Bin 7.3 reached at least suggestive significance in six analyses for ACR, serum creatinine concentration and creatinine clearance and Bin 16.4 reached suggestive significance in seven analyses for eGFR, serum creatinine concentration and creatinine clearance. Bins 1.1, 7.3, and 7.5 reached genome-wide significance in the subgroups of AI descent for eGFR, hypertension or diabetes for serum creatinine concentration and 5 for creatinine clearance in Chromosomes 1, 2, 3, 7, 9, 10, 11, 13, 14 and 16. Four bins, 7.3, 7.4, 10.4 and 16.4, reached suggestive significance for at least two phenotypes. Despite signals of genome-wide significance in subgroup analyses (Bins 1.1, 7.3 and 7.5), our main analyses do not pinpoint specific cytogenetic regions as linked with the examined renal phenotypes. The conducted genome scans did not agree beyond chance in the cytogenetic locations they prioritized.

**Comparison of results with findings of primary studies and with GWAS results**

Because no single bin attained genome-wide significance in the main analysis, we focused our attention on bins that reached suggestive significance for more than one phenotype (bins in Chromosomes 7, 10 and 16). Table 3 outlines previous genome scans for renal phenotypes with linkage findings which have reached genome-wide significance. Signals appear spread out across the genome although there are repeated signals on Chromosomes 2, 7 and 22 across phenotypes. Notably, apart from regions on Chromosome 7, most high LOD scores in Table 3 are in regions that are not selected in our meta-analysis, in large part attributable to heterogeneity across studies. For example, two studies found genome-wide significance for linkage to eGFR and serum creatinine on Chromosome 2 [29, 33], in the cytogenetic location corresponding to Bin 2.3. This bin was close to reaching suggestive significance in the meta-analysis (Figure 1), but with indications for high between-scan heterogeneity ($P$-value for heterogeneity $>0.90$).

Recent associations have been identified between CKD phenotypes and markers on Chromosome 16 [7] and Chromosome 22 [8–10, 38–40]. In our meta-analysis, we showed suggestive significance for eGFR in several subgroup analyses for bins 16.3 and 16.4, although none of the individual studies in Table 3 had a peak in Chromosome 16. On the other hand, two studies as shown in Table 3, reported linkage peaks in Chromosome 22 [41, 42] and we noted suggestive significance for bin 22.1 when data on eGFR, creatinine clearance and serum creatinine were combined. We did not note any other overlap with loci identified in recent meta-analysis of genome-wide association.
Fig. 2. GSMA for serum creatinine. Weighted average ranks from 11 population strata. Layout similar to Figure 1.

Fig. 3. GSMA for creatinine clearance. Weighted average ranks from four population strata. Layout similar to Figure 1.

Fig. 4. GSMA for ACR. Weighted average ranks from 11 population strata. Layout similar to Figure 1.
data for CKD [43, 44]. Kitsios and Zintzaras have investigated the agreement between genome-wide linkage scans and genome-wide association studies and noted ‘genomic convergence’ in 2 of 19 phenotypes. Although convergence could support true genetic effects, lack of congruence may reflect that these studies are designed to answer different questions (association studies are aimed at identifying common variants involved in disease etiology with small effects, whereas linkage approaches attempt to identify rare variants with larger effects) and employ different approaches. Perhaps the two types of research are better viewed as complementary in elucidating the genetics of common complex disease [45, 46].

**Issues affecting the results and their interpretation**

The genome scan meta-analysis has been applied to other complex phenotypes [21, 41, 42, 47–50] with the identification of several novel loci; loci with genome-wide significance have been identified in autism [21], hypertension [49] and inflammatory bowel disease and celiac disease [38, 50], schizophrenia [39], multiple sclerosis [40] and asthma and atopy [51]. However, diverse linkage signals across scans, possibly suggesting genetic heterogeneity across subsyndromes and subpopulations have been noted as a factor limiting effective synthesis of linkage data.

In our approach, we combined data across strata of different racial descent and health status. GSMA can identify regions that give weak but consistent linkage signals in multiple genome scans [42] but will not identify linked regions that are present in only a subset of scans for example because of population-specific effects [20]. Several studies included in the GSMA increased their statistical power by combining data across racial groups [28, 29, 33, 34]. While we employed a similar approach, we appreciate that combining across racial subgroups also increased heterogeneity. We therefore performed subgroup analyses by race and health status to explore subgroup-specific signals. In contrast to association studies, linkage analyses are family-based and therefore immune to population stratification. However, it is now appreciated that they may be affected by differences in LD patterns and range (admixture LD [52]) as well as allele frequencies across different ethnic groups which could potentially add to the heterogeneity of the linkage signal.

The synthesized outcomes in this study are surrogates of kidney disease. eGFR values depend on the estimating equation and factors such as race [53]; similarly, microalbuminuria can reverse [54] and may not uniformly correlate with severity of CKD; [55] treatment with inhibitors of the renin angiotensin system could minimize variability of both eGFR and ACR; finally, most studies have relied on single measurements which do not capture either the marked day to day variability of ACR or the change in eGFR over time [56]. Any of these factors may have impacted the results of the original study and accordingly the results of the meta-analysis.

Most included studies did not report unadjusted analyses, and therefore it is not possible to explore whether the negative findings of the meta-analysis are related to the choice of the analytic models in the primary studies. For example, if one adjusts for a factor that is in the path between the genetic information and the examined outcome, the relationship of the genetic information and the outcome could be ‘masked’. In studies which reported information from unadjusted analyses, the relative ranking across genetic markers was very similar to those from adjusted analyses.

An additional element of heterogeneity may have been introduced by combining original and digitized data. Digitizing is as best as accurate as the digitized figure. If the figure is of poor quality, digitizing may introduce errors.
compared to the primary data. In our case, almost all figures were of good quality. Limiting the analyses to the 13 strata from which we had primary data did not change the main findings. Nevertheless, it would be desirable that primary studies publish supplementary material with tabular data or with complete and good quality graphs.

Finally, we used bin sizes of ~30 cM. While this choice is somewhat arbitrary, it has some motivations from theory and simulation studies. The question of bin size for GSMA has been empirically explored by Hermanowski et al. [40] and discussed by Wise et al. [14] who recommended a bin width of ~30 cM, being wide enough to limit the correlation between adjacent bins and not too wide to avoid including distinct peak LOD scores from different studies within the same bin. Overall, the effect of different bin definitions is likely to be minor compared to other sources of error and heterogeneity in the genome-wide linkage studies analyzed [57].

Conclusion

To our knowledge, this is the first study to combine the findings from numerous existing genome linkage scans conducted for continuous traits of renal phenotypes. Overall, even after combining 22 genome scans, little evidence of consistent linkage on specific chromosomal regions was detected. While heterogeneity across populations, methodologies and study designs likely explains this lack of agreement, it is also possible that genome-wide scan methodologies have insufficient resolution for investigating complex renal phenotypes. Used in a complementary role to genome-wide association approaches, they may be valuable for uncovering private gene mutations that contribute to renal phenotypes within families.

Table 3. Observed results from genome scans with LOD scores >3.0a

<table>
<thead>
<tr>
<th>Author</th>
<th>Study</th>
<th>Phenotype</th>
<th>Type of population</th>
<th>Ethnicity</th>
<th>Genomic Location (cM)</th>
<th>LOD Score</th>
<th>GSMA Bin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr 1 Schelling [29]</td>
<td>FIND</td>
<td>GFR</td>
<td>Diabetic</td>
<td>Mexican-American</td>
<td>259</td>
<td>3.78</td>
<td>1.9</td>
</tr>
<tr>
<td>Chr 2 Freedman [33]</td>
<td>DHS</td>
<td>GFR</td>
<td>Diabetic/</td>
<td>African-American and</td>
<td>77</td>
<td>4.45</td>
<td>2.3</td>
</tr>
<tr>
<td>Schelling [29]</td>
<td>FIND</td>
<td>GFR</td>
<td>Diabetic</td>
<td>Mexican-American</td>
<td>82</td>
<td>3.02</td>
<td>2.4</td>
</tr>
<tr>
<td>Freedman [33]</td>
<td>DHS</td>
<td>creatinine Serum</td>
<td>Diabetic/</td>
<td>African-American and</td>
<td>80</td>
<td>3.55</td>
<td>2.3</td>
</tr>
<tr>
<td>Hunt [26]</td>
<td>Utah</td>
<td>creatinine Serum</td>
<td>General population</td>
<td>European-American</td>
<td>145</td>
<td>3.15</td>
<td>2.6</td>
</tr>
<tr>
<td>Freedmanb [36]</td>
<td></td>
<td></td>
<td>All-cause</td>
<td>African-American</td>
<td>186</td>
<td>3.05</td>
<td>2.7</td>
</tr>
<tr>
<td>Placha [37]</td>
<td>Joslin</td>
<td>Cystatin</td>
<td>Diabetic</td>
<td>European-American</td>
<td>202</td>
<td>4.1</td>
<td>2.8</td>
</tr>
<tr>
<td>Chr 3 Freedmanb [36]</td>
<td></td>
<td></td>
<td>All-cause</td>
<td>African-American</td>
<td>7</td>
<td>3.18</td>
<td>3.3</td>
</tr>
<tr>
<td>Chr 5 Krolewski [30]</td>
<td>Joslin</td>
<td>ACR</td>
<td>Diabetic</td>
<td>European-American</td>
<td>69</td>
<td>3.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Chr 6 Freedmanb [36]</td>
<td></td>
<td></td>
<td>All-cause</td>
<td>African-American</td>
<td>162.5</td>
<td>5.72</td>
<td>6.6</td>
</tr>
<tr>
<td>Chr 7 Turner [35]</td>
<td>GENOA</td>
<td>GFR</td>
<td>Hypertensive</td>
<td>African-American</td>
<td>43</td>
<td>3.65</td>
<td>7.2</td>
</tr>
<tr>
<td>Schelling [29]</td>
<td>FIND</td>
<td>GFR</td>
<td>Diabetic</td>
<td>Multiethnic</td>
<td>170</td>
<td>3.28</td>
<td>7.6</td>
</tr>
<tr>
<td>Schelling [29]</td>
<td>FIND</td>
<td>GFR</td>
<td>Diabetic</td>
<td>Mexican-American</td>
<td>170</td>
<td>4.23</td>
<td>7.6</td>
</tr>
<tr>
<td>Placha [37]</td>
<td>Joslin</td>
<td>Cystatin</td>
<td>Diabetic</td>
<td>European-American</td>
<td>23</td>
<td>4</td>
<td>7.1</td>
</tr>
<tr>
<td>Chr 8 Schelling [29]</td>
<td>FIND</td>
<td>GFR</td>
<td>Diabetic</td>
<td>Mexican-American</td>
<td>93</td>
<td>3.95</td>
<td>8.4</td>
</tr>
<tr>
<td>Chr 9 Arar [23]</td>
<td>SAFHS</td>
<td>GFR</td>
<td>Hypertensive</td>
<td>Mexican-American</td>
<td>75</td>
<td>3.87</td>
<td>9.3</td>
</tr>
<tr>
<td>Chr 10 Placha . [37]</td>
<td>Joslin</td>
<td>Cystatin</td>
<td>Diabetic</td>
<td>European-American</td>
<td>114</td>
<td>3.6</td>
<td>10.4</td>
</tr>
<tr>
<td>Chr 12 Mottl [24]</td>
<td>SHFS</td>
<td>GFR</td>
<td>General population</td>
<td>American Indian (Arizona)</td>
<td>39</td>
<td>3.5</td>
<td>12.2</td>
</tr>
<tr>
<td>Chr 13 Freedmanb [36]</td>
<td></td>
<td></td>
<td>All-cause</td>
<td>African-American</td>
<td>91</td>
<td>4.94</td>
<td>13.4</td>
</tr>
<tr>
<td>Chr 14 Leon [34]</td>
<td>HyperGEN</td>
<td>GFR</td>
<td>Hypertensive</td>
<td>African-American and</td>
<td>94</td>
<td>3.29</td>
<td>14.3</td>
</tr>
<tr>
<td>Chr 15 Leon [34]</td>
<td>HyperGEN</td>
<td>GFR</td>
<td>Hypertensive</td>
<td>African-American and</td>
<td>16</td>
<td>3.13</td>
<td>15.1</td>
</tr>
<tr>
<td>Chr 22 Krolewski [30]</td>
<td>Joslin</td>
<td>ACR</td>
<td>Diabetic</td>
<td>European-American</td>
<td>33</td>
<td>3.4</td>
<td>22.2</td>
</tr>
<tr>
<td>Pattaro [27]</td>
<td>Eurospan</td>
<td>Serum</td>
<td>General population</td>
<td>European-American</td>
<td>61</td>
<td>3.4</td>
<td>22.2</td>
</tr>
</tbody>
</table>

aDHS, Diabetes Heart Study; GFR, estimated glomerular filtration rate; FIND, Family Investigation of Nephropathy and Diabetes; GENOA, Genetic Epidemiology Network of Arteriopathy; HyperGEN, Hypertension Genetic Epidemiology Network Study; SAFHS, San Antonio Family Heart Study; SHFS, Strong Heart Family Study.
bStudies with dichotomous outcomes, not included in present analysis.
cNot analyzed as part of study.
dSubgroup in study.

Supplementary data

Supplementary data is available online at http://ndt.oxfordjournals.org/.

Acknowledgements. We gratefully acknowledge the help of Adrienne Williams, Julie Zeigler and Sarah Lalacci in collating and supplying data files and Dr Georgios Kitsios for providing information on which chromosomal locations correspond to each bin.

Funding. This work was supported by (i) National Institutes of Health (DK066992) to Dr Rao; (ii) Grant 1 UL1 RR025752-02, Tufts Clinical and...
Translational Science Institute; (iii) the National Heart, Lung and Blood Institute’s Framingham Heart Study (N01-HC-25195); (iv) Wake Forest analyses were supported by National Institutes of Health grants (R01 DK053591 to D.W.B.), (R01 HL56266 to B.I.F.), (R01 DK70941 to B.I.F.), (DK084149 to B.I.F.) and, in part, by the General Clinical Research Center of the Wake Forest University School of Medicine grant M01 RR07122; (v) HyperGEN study (HL54471, HL54472, HL54473, HL54495, HL54496, HL54497, HL54509, HL54515 and HL55673); (vi) (R01-A018734); (vii) The FIND study was supported by Research Grants (U01DK57292, U01DK57294, U01DK57295, U01DK57298, U01DK57300, U01DK57303, U01DK57304 and U01DK57299) from the NIDDK and, in part, by the Intramural Research Program of the NIDDK. This work was also supported by the National Center for Research Resources for the General Clinical Research Center grants: Case Western Reserve University (M01-RR-00080), Wake Forest University (M01-RR-07122), Harbor-UCLA Medical Center (M01-RR-00425), College of Medicine-University of California Irvine (M01-RR-002729), University of New Mexico ISC (M01-RR-00997) and Frederic C. Barter (M01-RR-01346).

Genotyping for the original studies was performed by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to the Johns Hopkins University, contract no. N01-HG-65403.

**Conflict of interest statement.** None declared.

**References**

Single nucleotidic polymorphism 844 A→G of FCAR is not associated with IgA nephropathy in Caucasians

Nicolas Maillard1,2,3,4, Lise Thibaudin5, Farida Abadjia1,2,3, Ingrid Masson1,2,3,4, Olivier Garraud1,2,3, François Berthoux4,5, Eric Alamartine1,2,3,4,5 and Christophe Mariat1,2,3,4,5

1EA3064, GIMAP, Université Jean Monnet, Saint Etienne France, 2PRES, Université de LYON, Lyon France, 3Université Jean Monnet, Saint Etienne France, 4Service de Néphrologie, Dialyse, Transplantation Rénale, CHU de Saint Etienne, Saint Etienne France and 5Laboratoire d’Explorations Fonctionnelles Rénales, CHU de Saint-Etienne, Saint Etienne France

Correspondence and offprint requests to: Nicolas Maillard; E-mail: nicolas.maillard@univ-st-etienne.fr

Abstract

Background. IgA nephropathy is characterized by a high heterogeneity of clinical expression with 10–30% of patients progressing to end-stage renal failure. The gene of the FcεRI or CD89 presents a single-nucleotide polymorphism responsible for a proinflammatory phenotype of neutrophils in vitro and ex vivo. The aim of our study was to assess whether this CD89 polymorphism 844 A→G is (i) a marker of disease susceptibility and/or (ii) associated with a more severe prognosis.

Methods. All patients diagnosed with IgA nephropathy and for whom DNA frozen sample was available were included in this European monocentric retrospective analysis and compared to a cohort of healthy volunteers. Allelic discrimination was performed by real-time quantitative polymerase chain reaction (Applied Biosystems™). We first compared the distribution of A and G alleles between patients and volunteers and then studied the relationships between alleles and renal function, risk factor for cardiovascular disease.