Rheumatoid arthritis (RA), like most other complex rheumatic diseases, has multiple genetic and environmental determinants contributing to its susceptibility. Evidence for the genetic contribution to RA can be seen in population-based and twin studies, as well as molecular investigations from candidate gene and linkage analysis. Despite extensive ongoing research, few candidate genes have been identified and validated in RA from large mixed populations. The lack of replication of initial associations may be a result of false-positive reportings (from multiple testing, population stratification, and genotyping errors), false-negative findings (due to small effect size coupled with inadequate sample size of validation studies), or population-specific differences. In this issue of The Journal, Oen, et al set out to characterize the RA phenotype, estimate familial incidence of RA, and evaluate the association of various cytokine genes implicated in the pathogenesis of RA in the North American Native (NAN) population from Manitoba and Northwest Ontario, Canada.

Homogenous populations, like the NAN, are now the focus of intensive interest for gene identification in complex diseases. Although the utility of such populations for this purpose remains unproven, it is widely assumed that homogenous populations offer advantages in gene identification by mitigating some of the above limitations. The benefits in pursuing gene identification studies in such populations include an increase in allele and locus homogeneity, as well as the often-discussed potential increase in linkage disequilibrium. The former is of importance given the etiologic heterogeneity that characterizes complex diseases. Thus the detection of a significant signal in RA from a mixed population is more challenging, as modest genetic signals may be overlooked in outbred populations, unless one assembles very large cohorts. Meanwhile, in homogenous populations there may be an increased signal to noise ratio to identify genes of modest risk, due to the relative genetic and environmental homogeneity that often exists within such populations.

We share Oen and colleagues’ enthusiasm in investigating the genetic determinants of RA in the NAN population, as there appears to be a greater genetic burden of RA in this population. This is based on the high prevalence of RA, an earlier age of onset, greater severity, as well as higher rate of seropositivity [including rheumatoid factor, shared epitope (SE), and antinuclear antibody] in the NAN RA population. Specifically, the Chipewa and Blackfoot Indians have a 5-fold increased rate of RA compared to the Caucasian North American and European populations (reviewed by Peschken and Esdaile). The ancestral history and migratory route of the NAN population likely has bearing on the increased prevalence of RA. The ancestors of NAN appeared to have originated from northeast Asia; patterns of migration of the NAN are likely to have created population bottlenecks, resulting in a small number of founding chromosomes. Thus the increased prevalence of RA in selected NAN populations is likely due to genetic drift, resulting from the change in allele frequencies associated with founder effects. For these reasons, the NAN population represents a unique resource for identification of RA related genes.

It should be noted, however, that within the NAN population there is likely to be heterogeneity due to multiple distinct waves of migration from different founders, together with later admixture between NAN groups. NAN subpopulations differ in age, size of the genetic bottleneck, and in expansion rate. Clinical heterogeneity can be seen, for example, by comparing Amerind Indians, who have increased rates of RA and connective tissue disease, with the Na-Dene Indians and Eskimos, who have high rates of spondyloarthropathies. The clinical heterogeneity among the NAN groups is mirrored in molecular studies. For instance, the specific SE alleles bearing association in the

See Familial seropositive RA in North American Native families: effects of shared epitope and cytokine genotypes, page 983
Tlingit, Yakima, and Pima (HLA-DRB1*1402) differ from those associated in the Chippewa (DRB1*04). Thus it is important to know as much as possible about the population structure and genealogical history to maximally utilize such populations for novel gene discoveries. Further, even within apparent homogenous populations, documenting self-reported ethnic background (for instance, the majority of the patients in Oen’s study were Cree or Ojibway and had 4 NAN grandparents) does not necessarily mean that no substructure exists within this group. This point is illustrated by a recent Icelandic study that investigated population substructure within this relatively homogenous genetic isolate. The investigators concluded that even in a homogenous population, various sampling strategies are required to take account of substructure, since there were small variations in allele frequencies by geographic region. Hence self-reported ethnicity is not sufficient for inferring the presence or absence of population substructure. Despite these caveats, we feel a more homogenous genetic background in a population will mean less molecular heterogeneity.

Oen’s report shows a higher familial prevalence of RA in the NAN population. Specifically, they noted that the prevalence of multiplex families among the RA probands whose families were studied was 50% (14 of 28 families); a lower bound for prevalence among relatives of all probands is therefore 17% (14/82). Although the high familial prevalence of RA is not unexpected given the epidemiological association with extreme caution, since apparent decreases in age at onset do not necessarily mean that no substructure exists within this group. This point is illustrated by a recent Icelandic study that investigated population substructure within this relatively homogenous genetic isolate. The investigators concluded that even in a homogenous population, various sampling strategies are required to take account of substructure, since there were small variations in allele frequencies by geographic region. Hence self-reported ethnicity is not sufficient for inferring the presence or absence of population substructure. Despite these caveats, we feel a more homogenous genetic background in a population will mean less molecular heterogeneity.

In summary, the NAN population is a valuable resource for identifying RA related genes, given the high genetic burden of this disease, coupled with the reduced allelic diversity compared with more outbred populations.

PROTON RAHMAN, MD, MSc, FRCP, Associate Professor of Medicine, Division of Rheumatology, Memorial University of Newfoundland, St. John’s, Newfoundland; CELIA GREENWOOD, Project Director, Program in Genetics and Genomic Biology, Hospital for Sick Children, Toronto, Ontario, Canada.

Address reprint requests to Dr. P. Rahman, St. Clare’s Mercy Hospital, 1 South – 154 LeMarchant Road, St. John’s, Newfoundland A1C 5B8. E-mail: prahman@mun.ca

REFERENCES