B. Cakir · J. S. Pankow · V. Salomaa · D. Couper ·
T. L. Morris · K. R. Brantley · K. M. Hiller · G. Heiss ·
B. W. Weston

Distribution of Lewis (FUT3) genotype and allele:
frequencies in a biethnic United States population

Received: 30 August 2001 / Accepted: 27 June 2002 / Published online: 11 October 2002
© Springer-Verlag 2002

Abstract The objective of the study was to examine the
prevalence and distribution of four major single nucleo-
tide polymorphisms (SNPs) (T59G, T1067G, T202C, and
C314T) of the Lewis (FUT3) gene in a biethnic United
States population. This population-based cross-sectional
study was based on data from the Atherosclerosis Risk in
Communities (ARIC) Study, which included 761 males
and females aged 45–64 years, who had no known/
detected clinical atherosclerotic disease (577 Caucasians,
184 African Americans). The main outcome measures
were prevalence of the Lewis genotype and allele
frequencies for four SNPs of the FUT3 gene. The most
common genotype was the “wild type” at all four
nucleotide positions (WWWW), which was found to be
present in 46.9% of ARIC participants. At least one
mutant allele was detected in 51.7% of Caucasians, and
56.7% of African Americans (P=0.59). The frequencies of
mutant alleles ranged from 6.3% to 18.4% at the four
FUT3 gene sites examined. The distribution of the Lewis
genotype and allele frequencies differed significantly by
ethnicity at sites 59, 202, and 314. The prevalence of the
Lewis genotype suggesting a lack of α(1,3/1,4)fucosyl-
transferase activity was 11.6% in Caucasians and 9.9% in
African Americans (P=0.67). Four specific SNPs of the
Lewis genotype are common in the population at large.
However, these four SNPs seem to fail to explain the
majority of Lewis-negative phenotype in African Amer-
icans, given that Lewis-negative genotype prevalence was
about one-third of what was expected. Use of rapid DNA
sequencing and simultaneous Lewis phenotype determi-
nation could avoid the problems associated with haplo-
type determination and Lewis genotype grouping. Further
studies testing SNPs of the Lewis gene are warranted, in
particular among African Americans.

Keywords Lewis gene · FUT3 gene ·
Human α(1,3/1,4)fucosyltransferase

Introduction

The Lewis antigen system is complex and not yet
completely understood. Lewis antigens are glycosphing-
golipids made by addition of monosaccharides onto
glycolipid and glycoprotein molecules [1]. Lewis antigens
and related structures [sialyl Lewis (a) and sialyl Lewis
(x)] have been shown to serve as components of cell
surface ligands for cell adhesion molecules [2], and on
tumor cells these antigens correlate with metastasizing
capacity [3, 4]. Lewis carbohydrate polymorphisms may
be of importance in selective survival under certain
environmental conditions, such as an attack by bacteria or
viruses [5]. Several studies have suggested an association
between the Lewis antigen system and various diseases
and/or conditions, including coronary heart disease
(CHD) and/or CHD-related risk factors [6, 7, 8, 9, 10,
11, 12, 13, 14, 15].

Most of the previous research has studied the Lewis
antigen system at a phenotypic level. Studies examining
the genetic background of Lewis phenotypes have been scarce, small in size, and limited to specific groups of individuals. The genetic control of Lewis antigen expression is complex because the Lewis gene (FUT3) and the secretor gene (FUT2) each encode a different fucosyltransferase, and the final oligosaccharide products are the result of the epistatic interactions of these enzymes on the same oligosaccharide receptors of type 1 [β-Gal(1→3)]β-GlcNAc-R and type 2 [β-Gal(1→4)]β-GlcNAc-R [1].

The FUT3 (Lewis) gene is polymorphic and located on chromosome 19. The Lewis gene behaves in a dominant fashion and codes for an α(1,3/1,4)fucosyltransferase that is expressed in exocrine secretions [16, 17]. The cDNA encoding Lewis fucosyltransferase has been isolated [16]. There are at least two alleles at the Lewis site: Le, which encodes a functional α(1,3/1,4)fucosyltransferase, and le, which is apparently silent. Approximately 95% of whites and 75% of blacks have at least one Le allele [17, 18]. Several alleles encode nonfunctional α(1,3/1,4)fucosyltransferases and are collectively known as le alleles. These silent variants contain single nucleotide polymorphisms (SNPs) in coding region of the FUT3 gene and are consistent with a decrease in enzyme levels and/or Lewis-negative phenotype in individuals [1]. Several single base mutations have been identified in a number of Lewis-negative individuals of various ethnic groups [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29].

The type and allele frequencies at the Lewis gene show striking differences in different populations and ethnic groups, and the Lewis genotype-phenotype correspondence is not straightforward [14, 25, 29]. Two previous studies have primarily investigated the Lewis genotype-phenotype correspondence in Caucasians [14, 25]. Salm et al. reported that between 90% and 95% of Lewis-negative individuals could be identified by screening for the SNPs, T59G, T1067A, T202C, C314T, and these four SNPs appear to be common in the population at large [14]. There is one study on FUT3 alleles in South African individuals, including the black Xhosa-Africans [29]. Pang et al. described five missense mutations C304A, T370G, G484A, G667A, and G808A in the Lewis gene and determined four novel alleles in the Xhosa population, two of which [i.e., FUT3(G484A,G667A) and FUT3(G484A,G667A,G808A)] are nonfunctional [29]. To our knowledge, no study has estimated the distribution of Lewis gene SNPs in African Americans.

We aimed to examine the prevalence of the four major single nucleotide polymorphisms (T59G, T1067A, T202C, and C314T) of the FUT3 gene in a biethnic United States population and to evaluate the distribution of Lewis genotype and allele frequencies by gender, ethnicity, and age. We also made an effort to infer phenotypic information based on the possible significance of the SNPs of the Lewis gene to investigate the correlation of these four mutations and Lewis-negative phenotype and to evaluate the potential utility of the molecular data alone for studying the association of the Lewis system and various disease(s)/condition(s).

### Materials and methods

Participants and methods

The study includes data from 761 individuals sampled from the Atherosclerosis Risk in Communities (ARIC) study. The ARIC cohort consists of a population-based probability sample of 15,792 individuals aged 45–64 years at the baseline survey, sampled from four United States communities: Forsyth County, North Carolina (includes the city of Winston-Salem), Jackson, Mississippi suburban Minneapolis, Minnesota, and Washington County, Maryland (includes the city of Hagerstown). The samples in Minneapolis and Washington county are primarily Caucasians and the sample in Forsyth includes about 15% African Americans. The Jackson center enrolled only African Americans. Detailed information on sampling schemes and other issues has been published elsewhere [30].

A stratified random sample of ARIC cohort members was selected to represent all African American and Caucasian ARIC participants at baseline exam (visit 1, 1986–1989) with the following exclusion criteria: (1) prevalent CHD (self-reported CHD history or ECG evidence) or missing data on prevalent CHD, (2) self-reported history of physician-diagnosed stroke or transient ischemic attack (TIA), (3) non-Caucasians in Minnesota and Washington County and participants in Forsyth County who were neither African American nor Caucasian were further excluded from the study population due to the small numbers of these groups (n=91 persons), and (4) six other individuals were excluded from analyses due to lack of strata information.

The eligible cohort members were stratified on age group (45–54, 55–64 years), gender, and cardio artery wall thickness (intima-media thickness, IMT) ≤ the 30th percentile or IMT >the 30th percentile). Within each stratum, selection was by simple random sampling.

Genotype information could be obtained for 761 (77.2% of the 986 individuals sampled) for all four sites of the Lewis gene (i.e., sites 59, 1067, 202, and 314) after excluding ARIC participants whose DNA samples could not be amplified after three attempts at genotyping.

Lewis genotyping

Genotyping was conducted at the University of North Carolina at Chapel Hill [14], using the frozen stored DNA samples obtained from the Central ARIC DNA laboratory in Houston, Texas, USA. Informed consent for the use of blood samples for DNA analyses for the aims of the ARIC study was obtained from all participants as part of the ARIC study.

Lewis genotyping in this study included the identification of four major mutations of the FUT3 gene at nucleotide positions 59, 1067, 202, and 314 (numbered according to the start codon at +73 in the original cloned cDNA) [16]. Of these, the T59G mutation exists in the transmembrane domain of the FUT3 gene and the other mutations are in the catalytic region [14, 31]. Allele-specific oligonucleotide hybridization (ASO) was performed as previously described [14].

The reliability of the Lewis genotyping was evaluated using replicate DNA samples from a randomly selected subgroup of 57 ARIC participants. Kappa coefficients with 95% confidence intervals were 0.92 (95% CI=0.75–1.00), 0.89 (95% CI=0.71–1.00), 0.97 (95% CI=0.91–1.00), and 0.87 (95% CI=0.75–0.99) for FUT3 gene SNPs at nucleotide positions 59, 1067, 202, and 314, respectively. The number of discordant pairs was one for sites 59, 1067, and 202 and four for site 314, corresponding to percent agreement estimates of 93% or higher at each of the four sites. This indicated that the Lewis genotyping by allele-specific oligonucleotide hybridization method is a highly reliable procedure.