Supplementary Online Content

Auer PL, Nalls M, Meschia JF, et al; National Heart, Lung, and Blood Institute Exome Sequencing Project. Rare and coding region genetic variants associated with risk of ischemic stroke: the NHLBI Exome Sequence Project. *JAMA Neurol*. Published online May 11, 2015. doi:10.1001/jamaneurol.2015.0582.

eAppendix 1. Membership of the NHBLI Exome Sequencing Project (ESP)

eAppendix 2. Exome Sequencing and Variant Calling

eAppendix 3. ExomeChip Genotyping and Quality Control

eAppendix 4. Statistical Analyses

eFigure 1. Principal Components 1 (PC1) and 2 (PC2) From the ESP6800 With Self-reported Ancestry

eFigure 2. Density Distribution of LOD Scores per Subset Of IBD Tests in Affected Sib Pairs With Ischemic Stroke

eTable 1. Meta-analysis of Single-Variant and Gene-Based Association Results With Ischemic Stroke for Previously Published Candidate Genes/Loci

eTable 2. Meta-analysis of Single-Variant and Gene-Based Association Results With Ischemic Stroke and Small Vessel Subtype for Previously Published Candidate Genes/Loci

eTable 3. Meta-analysis of Single-Variant and Gene-Based Association Results With Ischemic Stroke and Large Vessel Subtype for Previously Published Candidate Genes/Loci

eTable 4. Meta-analysis of Single-Variant and Gene-Based Association Results With Ischemic Stroke and Cardioembolic Subtype for Previously Published Candidate Genes/Loci

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix 4. Membership of the NHBLI Exome Sequencing Project (ESP)

BroadGO

Stacey B. Gabriel (Broad Institute)[4, 5, 11, 16, 17], David M. Altshuler (Broad Institute, Harvard Medical School, Massachusetts General Hospital)[1, 5, 7, 17], Gonçalo R. Abecasis (University of Michigan)[3, 5, 9, 13, 15, 17], Hooman Allayee (University of Southern California)[5], Sharon Cresci (Washington University School of Medicine)[5], Mark J. Daly (Broad Institute, Massachusetts General Hospital), Paul I. W. de Bakker (Broad Institute, Harvard Medical School, University Medical Center Utrecht)[3, 15], Mark A. DePristo (Broad Institute)[4, 13, 15, 16], Ron Do (Broad Institute)[5, 9, 13, 15], Peter Donnelly (University of Oxford)[5], Deborah N. Farlow (Broad Institute)[3, 4, 5, 12, 14, 16, 17], Tim Fennell (Broad Institute), Kiran Garimella (University of Oxford)[4, 16], Stanley L. Hazen (Cleveland Clinic)[5], Youna Hu (University of Michigan)[3, 9, 15], Daniel M. Jordan (Harvard Medical School, Harvard University)[13], Goo Jun (University of Michigan)[13], Sekar Kathiresan (Broad Institute, Harvard Medical School, Massachusetts General Hospital)[5, 8, 9, 12, 14, 15, 17, 20], Hyun Min Kang (University of Michigan)[9, 13, 16], Adam Kiezun (Broad Institute)[5, 13, 15], Guillaume Lettre (Broad Institute, Montreal Heart Institute, Université de Montréal)[1, 2, 13, 15], Bingshan Li (University of Michigan)[3], Mingyao Li (University of Pennsylvania)[5], Christopher H. Newton-Cheh (Broad Institute, Massachusetts General Hospital, Harvard Medical School)[3, 8, 15], Sandosh Padmanabhan (University of Glasgow School of Medicine)[3, 12, 15], Gina Peloso (Broad Institute, Harvard Medical School, Massachusetts General Hospital)[5], Sara Pulit (Broad Institute)[3, 15], Daniel J. Rader (University of Pennsylvania)[5], David Reich (Broad Institute, Harvard Medical School)[15], Muredach P. Reilly (University of Pennsylvania)[5], Manuel A. Rivas (Broad Institute, Massachusetts General Hospital)[5], Steve Schwartz (Fred Hutchinson Cancer Research Center)[5, 12], Laura Scott (University of Michigan)[1], David S. Siscovick (University of Washington)[5, 1, 25], John A. Spertus (University of Missouri Kansas City)[5], Nathan O. Stitziel (Brigham and Women's Hospital)[5, 15], Nina Stoletzki (Brigham and Women's Hospital, Broad Institute, Harvard Medical School)[13], Shamil R. Sunyaev (Brigham and Women's Hospital, Broad Institute, Harvard Medical School)[1, 3, 5, 13, 15], Benjamin F. Voight (Broad Institute, Massachusetts General Hospital), Cristen J. Willer (University of Michigan)[1, 9, 13, 15]

HeartGO

Stephen S. Rich (University of Virginia)[2, 4, 7, 8, 9, 11,14, 15, 17, 18, 31], Ermeg Akylbekova (Jackson State University, University of Mississippi Medical Center)[29], Larry D. Atwood* (Boston University)[1, 11, 28], Christie M. Ballantyne (Baylor College of Medicine, Methodist DeBakey Heart Center)[9, 22], Maja Barbalic (University of Texas Health Science Center Houston)[9, 14, 15, 17, 22], R. Graham Barr (Columbia University Medical Center)[10, 31], Emelia J. Benjamin (Boston University)[14, 20, 28], Joshua Bis (University of Washington)[15, 23], Eric Boerwinkle (University of Texas Health Science Center Houston)[3, 5, 9, 13, 15, 17, 22], Donald W. Bowden (Wake Forest University)[1, 31], Jennifer Brody (University of Washington)[3, 5, 15, 23], Matthew Budoff (Harbor-UCLA Medical Center)[31], Greg Burke (Wake Forest University)[5, 31], Sarah Buxbaum (Jackson State University)[3, 13, 15, 29], Jeff Carr (Wake Forest University)[25, 29, 31], Donna T. Chen (University of Virginia)[6, 11], Ida Y. Chen (Cedars-Sinai Medical Center)[1, 31], Wei-Min Chen (University of Virginia)[13, 15, 18], Pat Concannon (University of Virginia)[11], Jacy Crosby (University of Texas Health Science Center Houston)[22], L. Adrienne Cupples (Boston University)[1, 3, 5, 9, 13, 15, 18, 28], Ralph D'Agostino (Boston University)[28], Anita L. DeStefano (Boston University)[13, 18, 28], Albert Dreisbach (University of Mississippi Medical Center)[3, 29], Josée Dupuis (Boston University)[1, 28], J. Peter Durda (University of Vermont)[15, 23], Jaclyn Ellis (University of North Carolina Chapel Hill)[1], Aaron R. Folsom (University of Minnesota)[5, 22], Myriam Fornage (University of Texas Health Science Center Houston)[3, 18, 25], Caroline S. Fox (National Heart, Lung, and Blood Institute)[1, 28], Ervin Fox (University of Mississippi Medical Center)[3, 9, 29], Vincent Funari (Cedars-Sinai Medical Center)[1, 11, 31], Santhi K. Ganesh (University of Michigan)[2, 22], Julius Gardin (Hackensack University Medical Center)[25], David Goff (Wake Forest University)[25], Ora Gordon (Cedars-Sinai Medical Center)[11, 31], Wayne Grody (University of California Los Angeles)[11, 31], Myron Gross (University of Minnesota)[1, 5, 14, 25], Xiuqing Guo (Cedars-Sinai Medical Center)[3, 15, 31], Ira M. Hall (University of Virginia), Nancy L. Heard-Costa (Boston University)[1, 11, 28], Susan R. Heckbert (University of Washington)[10, 14, 20, 23], Nicholas Heintz (University of Vermont), David M. Herrington (Wake Forest University)[5, 31], DeMarc Hickson (Jackson State University, University of Mississippi Medical Center)[29], Jie Huang (National Heart, Lung, and Blood Institute)[5, 28], Shih-Jen Hwang (Boston

University, National Heart, Lung, and Blood Institute)[3, 28], David R. Jacobs (University of Minnesota)[25], Nancy S. Jenny (University of Vermont)[1, 2, 23], Andrew D. Johnson (National Heart, Lung, and Blood Institute)[2, 5, 11, 28], Craig W. Johnson (University of Washington)[15, 31], Steven Kawut (University of Pennsylvania)[10,31], Richard Kronmal (University of Washington)[31], Raluca Kurz (Cedars-Sinai Medical Center)[11, 31], Ethan M. Lange (University of North Carolina Chapel Hill)[3, 5, 9, 13, 34], Leslie A. Lange (University of North Carolina Chapel Hill)[1, 2, 3, 5, 9, 12, 13, 15, 17, 18, 20, 25, 34], Martin G. Larson (Boston University)[3, 15, 28], Mark Lawson (University of Virginia), Cora E. Lewis (University of Alabama at Birmingham)[25,34], Daniel Levy (National Heart, Lung, and Blood Institute)[3, 15, 17, 28], Dalin Li (Cedars-Sinai Medical Center)[11, 15, 31], Honghuang Lin (Boston University)[20, 28], Chunyu Liu (National Heart, Lung, and Blood Institute)[3, 28], Jiankang Liu (University of Mississippi Medical Center)[1, 29], Kiang Liu (Northwestern University)[25], Xiaoming Liu (University of Texas Health Science Center Houston)[15, 22], Yongmei Liu (Wake Forest University)[2, 5, 31], W. T. Longstreth, Jr. (University of Washington)[18, 23], Cay Loria (National Heart, Lung, and Blood Institute)[25], Thomas Lumley (University of Auckland)[9, 23], Kathryn Lunetta (Boston University)[28], Aaron J. Mackey (University of Virginia)[16, 18], Rachel Mackey (University of Pittsburgh)[1, 23, 31], Ani Manichaikul (University of Virginia)[8, 15, 18, 31], Taylor Maxwell (University of Texas Health Science Center Houston)[22], Barbara McKnight (University of Washington)[15, 23], James B. Meigs (Brigham and Women's Hospital, Harvard Medical School, Massachusetts General Hospital)[1, 28], Alanna C. Morrison (University of Texas Health Science Center Houston)[3, 15, 17], Solomon K. Musani (University of Mississippi Medical Center)[3, 29], Josyf C. Mychaleckyj (University of Virginia)[13, 15, 31], Jennifer A. Nettleton (University of Texas Health Science Center Houston)[9, 22], Kari North (University of North Carolina Chapel Hill)[1, 3, 9, 10, 13, 15, 17, 34], Christopher J. O'Donnell (Massachusetts General Hospital, National Heart, Lung, and Blood Institute)[2, 5, 9, 11, 12, 14, 15, 17, 20, 28], Daniel O'Leary (Tufts University School of Medicine)[25, 31], Frank S. Ong (Cedars-Sinai Medical Center)[3, 11, 31], Walter Palmas (Columbia University)[3, 15, 31], James S. Pankow (University of Minnesota)[1, 22], Nathan D. Pankratz (Indiana University School of Medicine)[15, 25], Shom Paul (University of Virginia), Marco Perez (Stanford University School of Medicine), Sharina D. Person (University of Alabama at Birmingham, University of Alabama at Tuscaloosa)[25], Joseph Polak (Tufts University School of Medicine)[31], Wendy S. Post (Johns Hopkins University)[3, 9, 11, 14, 20, 31], Bruce M. Psaty (Group Health Research Institute, University of Washington)[3, 5, 9, 11, 14, 15, 23], Aaron R. Quinlan (University of Virginia)[18, 19], Leslie J. Raffel (Cedars-Sinai Medical Center)[6, 11, 31], Vasan S. Ramachandran (Boston University)[3, 28], Alexander P. Reiner (Fred Hutchinson Cancer Research Center, University of Washington)[1, 2, 3, 5, 9, 11, 12, 13, 14, 15, 20, 25, 34], Kenneth Rice (University of Washington)[15, 23], Jerome I. Rotter (Cedars-Sinai Medical Center)[1, 3, 6, 8, 11, 15, 31], Jill P. Sanders (University of Vermont)[23], Pamela Schreiner (University of Minnesota)[25], Sudha Seshadri (Boston University)[18, 28], Steve Shea (Brigham and Women's Hospital, Harvard University)[28], Stephen Sidney (Kaiser Permanente Division of Research, Oakland, CA)[25], Kevin Silverstein (University of Minnesota)[25], David S. Siscovick (University of Washington)[5, 1, 25], Nicholas L. Smith (University of Washington)[2, 15, 20, 23], Nona Sotoodehnia (University of Washington)[3, 15, 23], Asoke Srinivasan (Tougaloo College)[29], Herman A. Taylor (Jackson State University, Tougaloo College, University of Mississippi Medical Center)[5, 29], Kent Taylor (Cedars-Sinai Medical Center)[31], Fridtjof Thomas (University of Texas Health Science Center Houston)[3, 22], Russell P. Tracy (University of Vermont)[5, 9, 11, 12, 14, 15, 17, 20, 23], Michael Y. Tsai (University of Minnesota)[9, 31], Kelly A. Volcik (University of Texas Health Science Center Houston)[22], Chrstina L Wassel (University of California San Diego)[9, 15, 31], Karol Watson (University of California Los Angeles)[31], Gina Wei (National Heart, Lung, and Blood Institute)[25], Wendy White (Tougaloo College)[29], Kerri L. Wiggins (University of Vermont)[23], Jemma B. Wilk (Boston University)[10, 28], O. Dale Williams (Florida International University)[25], Gregory Wilson (Jackson State University)[29], James G. Wilson (University of Mississippi Medical Center)[1, 2, 5, 8, 9, 11, 12, 14, 17, 20, 29], Phillip Wolf (Boston University)[28], Neil A. Zakai (University of Vermont)[2, 23]

ISGS and SWISS

John Hardy (Reta Lila Weston Research Laboratories, Institute of Neurology, University College London)[18], James F. Meschia (Mayo Clinic)[18], Michael Nalls (National Institute on Aging)[2, 18], Stephen S. Rich (University of Virginia)[2, 4, 7, 8, 9, 11, 14, 15, 17, 18, 31], Andrew Singleton (National Institute on Aging)[18], Brad Worrall (University of Virginia)[18]

Michael J. Bamshad (Seattle Children's Hospital, University of Washington)[4, 6, 7, 8, 10, 11, 13, 15, 17, 27], Kathleen C. Barnes (Johns Hopkins University)[2, 10, 12, 14, 15, 17, 20, 24, 30, 32], Ibrahim Abdulhamid (Children's Hospital of Michigan)[27], Frank Accurso (University of Colorado)[27], Ran Anbar (Upstate Medical University)[27], Terri Beaty (Johns Hopkins University)[24, 30], Abigail Bigham (University of Washington)[13, 15, 27], Phillip Black (Children's Mercy Hospital)[27], Eugene Bleecker (Wake Forest University)[33], Kati Buckingham (University of Washington)[27], Anne Marie Cairns (Maine Medical Center)[27], Wei-Min Chen (University of Virginia)[13, 15, 18], Daniel Caplan (Emory University)[27], Barbara Chatfield (University of Utah)[27], Aaron Chidekel (A.I. Dupont Institute Medical Center)[27], Michael Cho (Brigham and Women's Hospital, Harvard Medical School)[13, 15, 24], David C. Christiani (Massachusetts General Hospital)[21], James D. Crapo (National Jewish Health)[24, 30], Julia Crouch (Seattle Children's Hospital)6, Denise Daley (University of British Columbia)[30], Anthony Dang (University of North Carolina Chapel Hill)[26], Hong Dang (University of North Carolina Chapel Hill)[26], Alicia De Paula (Ochsner Health System)[27], Joan DeCelie-Germana (Schneider Children's Hospital)[27], Allen Dozor (New York Medical College, Westchester Medical Center)[27], Mitch Drumm (University of North Carolina Chapel Hill)[26], Maynard Dyson (Cook Children's Med. Center)[27], Julia Emerson (Seattle Children's Hospital, University of Washington)[27], Mary J. Emond (University of Washington)[10, 13, 15, 17, 27], Thomas Ferkol (St. Louis Children's Hospital, Washington University School of Medicine)[27], Robert Fink (Children's Medical Center of Dayton)[27], Cassandra Foster (Johns Hopkins University)[30], Deborah Froh (University of Virginia)[27], Li Gao (Johns Hopkins University)[24, 30, 32], William Gershan (Children's Hospital of Wisconsin)[27], Ronald L. Gibson (Seattle Children's Hospital, University of Washington)[10, 27], Elizabeth Godwin (University of North Carolina Chapel Hill)[26], Magdalen Gondor (All Children's Hospital Cystic Fibrosis Center)[27], Hector Gutierrez (University of Alabama at Birmingham)[27], Nadia N. Hansel (Johns Hopkins University, Johns Hopkins University School of Public Health)[10, 15, 30], Paul M. Hassoun (Johns Hopkins University)[10, 14, 32], Peter Hiatt (Texas Children's Hospital)[27], John E. Hokanson (University of Colorado)[24], Michelle Howenstine (Indiana University, Riley Hospital for Children)[27], Laura K. Hummer (Johns Hopkins University)[32], Seema M. Jamal (University of Washington)[11], Jamshed Kanga (University of Kentucky)[27], Yoonhee Kim (National Human Genome Research Institute)[24, 32], Michael R. Knowles (University of North Carolina Chapel Hill)[10, 26], Michael Konstan (Rainbow Babies & Children's Hospital)[27], Thomas Lahiri (Vermont Children's Hospital at Fletcher Allen Health Care)[27], Nan Laird (Harvard School of Public Health)[24], Christoph Lange (Harvard School of Public Health)[24], Lin Lin (Harvard Medical School)[21], Xihong Lin (Harvard School of Public Health)[21], Tin L. Louie (University of Washington)[13, 15, 27], David Lynch (National Jewish Health)[24], Barry Make (National Jewish Health)[24], Thomas R. Martin (University of Washington, VA Puget Sound Medical Center)[10, 21], Steve C. Mathai (Johns Hopkins University)[32], Rasika A. Mathias (Johns Hopkins University)[10, 13, 15, 30, 32], John McNamara (Children's Hospitals and Clinics of Minnesota)[27], Sharon McNamara (Seattle Children's Hospital)[27], Deborah Meyers (Wake Forest University)[33], Susan Millard (DeVos Children's Butterworth Hospital, Spectrum Health Systems)[27], Peter Mogayzel (Johns Hopkins University)[27], Richard Moss (Stanford University)[27], Tanda Murray (Johns Hopkins University)[30], Dennis Nielson (University of California at San Francisco)[27], Blakeslee Noyes (Cardinal Glennon Children's Hospital)[27], Wanda O'Neal (University of North Carolina Chapel Hill)[26], David Orenstein (Children's Hospital of Pittsburgh)[27], Brian O'Sullivan (University of Massachusetts Memorial Health Care)[27], Rhonda Pace (University of North Carolina Chapel Hill)[26], Peter Pare (St. Paul's Hospital)[30], H. Worth Parker (Dartmouth-Hitchcock Medical Center, New Hampshire Cystic Fibrosis Center)[27], Mary Ann Passero (Rhode Island Hospital)[27], Elizabeth Perkett (Vanderbilt University)[27], Adrienne Prestridge (Children's Memorial Hospital)[27], Nicholas M. Rafaels (Johns Hopkins University)[30], Bonnie Ramsey (Seattle Children's Hospital, University of Washington)[27], Elizabeth Regan (National Jewish Health)[24], Clement Ren (University of Rochester)[27], George Retsch-Bogart (University of North Carolina Chapel Hill)[27], Michael Rock (University of Wisconsin Hospital and Clinics)[27], Antony Rosen (Johns Hopkins University)[32], Margaret Rosenfeld (Seattle Children's Hospital, University of Washington)[27], Ingo Ruczinski (Johns Hopkins University School of Public Health)[13, 15, 30], Andrew Sanford (University of British Columbia)[30], David Schaeffer (Nemours Children's Clinic)[27], Cindy Sell (University of North Carolina Chapel Hill)[26], Daniel Sheehan (Children's Hospital of Buffalo)[27], Edwin K. Silverman (Brigham and Women's Hospital, Harvard Medical School)[24, 30], Don Sin (Children's Medical Center of Dayton)[30], Terry Spencer (Elliot Health System)[27], Jackie Stonebraker (University of North Carolina Chapel Hill)[26], Holly K. Tabor (Seattle Children's Hospital, University of Washington)[6, 10, 11, 17, 27], Laurie Varlotta (St.

Christopher's Hospital for Children)[27], Candelaria I. Vergara (Johns Hopkins University)[30], Robert Weiss [30], Fred Wigley (Johns Hopkins University)[32], Robert A. Wise (Johns Hopkins University)[30], Fred A. Wright (University of North Carolina Chapel Hill)[26], Mark M. Wurfel (University of Washington)[10, 14, 21], Robert Zanni (Monmouth Medical Center)[27], Fei Zou (University of North Carolina Chapel Hill)[26]

SeattleGO

Deborah A. Nickerson (University of Washington)[3, 4, 5, 7, 8, 9, 11, 15, 17, 18, 19], Mark J. Rieder (University of Washington)[4, 11, 13, 15, 16, 17, 19], Phil Green (University of Washington), Jay Shendure (University of Washington)[1, 8, 14, 16, 17], Joshua M. Akey (University of Washington)[13, 14, 15], Michael J. Bamshad (Seattle Children's Hospital, University of Washington)[4, 6, 7, 8, 10, 11, 13, 15, 17, 27], Kristine L. Bucasas (Baylor College of Medicine)[15], Carlos D. Bustamante (Stanford University School of Medicine)[3, 13, 15], David R. Crosslin (University of Washington)[2, 9], Evan E. Eichler (University of Washington)[19], P. Keolu Fox[2], Wenqing Fu (University of Washington)[13], Adam Gordon (University of Washington)[11], Simon Gravel (Stanford University School of Medicine)[13, 15], Gail P. Jarvik (University of Washington)[9, 15], Jill M. Johnsen (Puget Sound Blood Center, University of Washington)[2], Mengyuan Kan (Baylor College of Medicine)[13], Eimear E. Kenny (Stanford University School of Medicine)[3, 13, 15], Jeffrey M. Kidd (Stanford University School of Medicine)[13, 15], Fremiet Lara-Garduno (Baylor College of Medicine)[15], Suzanne M. Leal (Baylor College of Medicine)[1, 13, 15, 16, 17, 19, 20], Dajiang J. Liu (Baylor College of Medicine)[13, 15], Sean McGee (University of Washington)[13, 15, 19], Timothy D. O'Connor (University of Washington)[13], Bryan Paeper (University of Washington)[16], Peggy D. Robertson (University of Washington)[4], Joshua D. Smith (University of Washington)[4, 16, 19], Jeffrey C. Staples (University of Washington), Jacob A. Tennessen (University of Washington)[13], Emily H. Turner (University of Washington)[4, 16], Gao Wang (Baylor College of Medicine)[1,13,20], Qian Yi (University of Washington)[4]

WHISP

Rebecca Jackson (Ohio State University)[1, 2, 4, 5, 8, 12, 14, 15, 17, 18, 20, 34], Kari North (University of North Carolina Chapel Hill)[1, 3, 9, 10, 13, 15, 17, 34], Ulrike Peters (Fred Hutchinson Cancer Research Center)[1, 3, 11, 12, 13, 15, 17, 18, 34], Christopher S. Carlson (Fred Hutchinson Cancer Research Center, University of Washington)[1, 2, 3, 5, 12, 13, 14, 15, 16, 17, 18, 19, 34], Garnet Anderson (Fred Hutchinson Cancer Research Center)[34], Hoda Anton-Culver (University of California at Irvine)[34], Themistocles L. Assimes (Stanford University School of Medicine)[5, 9, 11, 34], Paul L. Auer (Fred Hutchinson Cancer Research Center)[1, 2, 3, 5, 11, 12, 13, 15, 16, 18, 34], Shirley Beresford (Fred Hutchinson Cancer Research Center)[34], Chris Bizon (University of North Carolina Chapel Hill)[3, 9, 13, 15, 34], Henry Black (Rush Medical Center)[34], Robert Brunner (University of Nevada)[34], Robert Brzyski (University of Texas Health Science Center San Antonio)[34], Dale Burwen (National Heart, Lung, and Blood Institute WHI Project Office)[34], Bette Caan (Kaiser Permanente Division of Research, Oakland, CA)[34], Cara L. Carty (Fred Hutchinson Cancer Research Center)[18, 34], Rowan Chlebowski (Los Angeles Biomedical Research Institute)[34], Steven Cummings (University of California at San Francisco)[34], J. David Curb* (University of Hawaii)[9, 18, 34], Charles B. Eaton (Brown University, Memorial Hospital of Rhode Island)[12, 34], Leslie Ford (National Heart, Lung, and Blood Institute, National Heart, Lung, and Blood Institute WHI Project Office)[34], Nora Franceschini (University of North Carolina Chapel Hill)[2, 3, 9, 10, 15, 34], Stephanie M. Fullerton (University of Washington)[6, 11, 34], Margery Gass (University of Cincinnati)[34], Nancy Geller (National Heart, Lung, and Blood Institute WHI Project Office)[34], Gerardo Heiss (University of North Carolina Chapel Hill)[5, 34], Barbara V. Howard (Howard University, MedStar Research Institute)34, Li Hsu (Fred Hutchinson Cancer Research Center)[1, 13, 15, 18, 34], Carolyn M. Hutter (Fred Hutchinson Cancer Research Center)[13, 15, 18, 34], John Ioannidis (Stanford University School of Medicine)[11, 34], Shuo Jiao (Fred Hutchinson Cancer Research Center)[34], Karen C. Johnson (University of Tennessee Health Science Center)[3, 34], Charles Kooperberg (Fred Hutchinson Cancer Research Center)[1, 5, 9, 13, 14, 15, 17, 18, 34], Lewis Kuller (University of Pittsburgh)[34], Andrea LaCroix (Fred Hutchinson Cancer Research Center)[34], Kamakshi Lakshminarayan (University of Minnesota)[18, 34], Dorothy Lane (State University of New York at Stony Brook)[34], Ethan M. Lange (University of North Carolina Chapel Hill)[3, 5, 9, 13, 34], Leslie A. Lange (University of North Carolina Chapel Hill)[1, 2, 3, 5, 9, 12, 13, 15, 17, 18, 20, 25, 34], Norman Lasser (University of Medicine and Dentistry of New Jersey)[34], Erin LeBlanc (Kaiser Permanente Center for Health Research, Portland, OR)[34], Cora E. Lewis (University of Alabama at Birmingham)[25,34], Kuo-Ping Li (University of North Carolina Chapel Hill)[9, 34], Marian Limacher (University of Florida)[34], Dan-Yu Lin

(University of North Carolina Chapel Hill)[1, 3, 9, 13, 15, 34], Benjamin A. Logsdon (Fred Hutchinson Cancer Research Center)[2, 34], Shari Ludlam (National Heart, Lung, and Blood Institute WHI Project Office)[34], JoAnn E. Manson (Brigham and Women's Hospital, Harvard School of Public Health)[34], Karen Margolis (University of Minnesota)[34], Lisa Martin (George Washington University Medical Center)[9, 34], Joan McGowan (National Heart, Lung, and Blood Institute WHI Project Office)[34], Keri L. Monda (Amgen, Inc.)[1, 15, 34], Jane Morley Kotchen (Medical College of Wisconsin)[34], Lauren Nathan (University of California Los Angeles)[34], Judith Ockene (Fallon Clinic, University of Massachusetts)[34], Mary Jo O'Sullivan (University of Miami)[34], Lawrence S. Phillips (Emory University)[34], Ross L. Prentice (Fred Hutchinson Cancer Research Center)[34], Alexander P. Reiner (Fred Hutchinson Cancer Research Center, University of Washington)[1, 2, 3, 5, 9, 11, 12, 13, 14, 15, 20, 25, 34], John Robbins (University of California at Davis)[34], Jennifer G. Robinson (University of Iowa)[9, 11, 18, 34], Jacques E. Rossouw (National Heart, Lung, and Blood Institute, National Heart, Lung, and Blood Institute WHI Project Office)[5, 14, 17, 20, 34], Haleh Sangi-Haghpeykar (Baylor College of Medicine)[34], Gloria E. Sarto (University of Wisconsin)[34], Sally Shumaker (Wake Forest University)[34], Michael S. Simon (Wayne State University)[34], Marcia L. Stefanick (Stanford University School of Medicine)[34], Evan Stein (Medical Research Labs)[34], Hua Tang (Stanford University)[2, 34], Kira C. Taylor (University of Louisville)[1, 3, 13, 15, 20, 34], Cynthia A. Thomson (University of Arizona)[34], Timothy A. Thornton (University of Washington)[13, 15, 18, 34], Linda Van Horn (Northwestern University)[34], Mara Vitolins (Wake Forest University)[34], Jean Wactawski-Wende (University of Buffalo)[34], Robert Wallace (University of Iowa)[2, 34], Sylvia Wassertheil-Smoller (Boston University)[18, 34], Donglin Zeng (University of North Carolina Chapel Hill)[9, 34]

*deceased

NHLBI GO ESP Project Team

Deborah Applebaum-Bowden (National Heart, Lung, and Blood Institute)[4, 7, 12, 17], Michael Feolo (National Center for Biotechnology Information)[12], Weiniu Gan (National Heart, Lung, and Blood Institute)[7, 8, 16, 17], Dina N. Paltoo (National Heart, Lung, and Blood Institute)[4, 6, 11, 17], Jacques E. Rossouw (National Heart, Lung, and Blood Institute, National Heart, Lung, and Blood Institute WHI Project Office)[5, 14, 17, 20, 34], Phyliss Sholinsky (National Heart, Lung, and Blood Institute)[4, 12, 17], Anne Sturcke (National Center for Biotechnology Information)[12]

ESP Groups

[1]Anthropometry Project Team, [2]Blood Count/Hematology Project Team, [3]Blood Pressure Project Team, [4]Data Flow Working Group, [5]Early MI Project Team, [6]ELSI Working Group, [7]Executive Committee, [8]Family Study Project Team, [9]Lipids Project Team, [10]Lung Project Team, [11]Personal Genomics Project Team, [12]Phenotype and Harmonization Working Group, [13]Population Genetics and Statistical Analysis Working Group, [14]Publications and Presentations Working Group, [15]Quantitative Analysis Ad Hoc Task Group, [16]Sequencing and Genotyping Working Group, [17]Steering Committee, [18]Stroke Project Team, [19]Structural Variation Working Group, [20]Subclinical/Quantitative Project Team

ESP Cohorts

[21]Acute Lung Injury (ALI), [22]Atherosclerosis Risk in Communities (ARIC), [23]Cardiovascular Health Study (CHS), [24]Chronic Obstructive Pulmonary Disease (COPDGene), [25]Coronary Artery Risk Development in Young Adults (CARDIA), [26]Cystic Fibrosis (CF), [27]Early Pseudomonas Infection Control (EPIC), [28]Framingham Heart Study (FHS), [29]Jackson Heart Study (JHS), [30]Lung Health Study (LHS), [31]Multi-Ethnic Study of Atherosclerosis (MESA), [32]Pulmonary Arterial Hypertension (PAH), [33]Severe Asthma Research Program (SARP), [34]Women's Health Initiative (WHI)

eAppendix 2. Exome Sequencing and Variant Calling

Exome Sequencing

Exome sequencing was performed at the Broad Institute of the Massachusetts Institute of Technology and Harvard University (Broad) and at the University of Washington (UW). The processes of library construction, exome capture, sequencing, and mapping were performed as previously described.[1] Samples of DNA were quality controlled by concentration estimation by Pico Green and in some cases by gel electrophoresis and real-time PCR-based genotyping. For the majority of the samples, other than those from the WHI, initial QC was performed centrally at the University of Vermont prior to shipping to the UW and the Broad.

In both centers, DNA samples were prepared by subjecting genomic DNA to shearing followed by ligation of sequencing adaptors. Exome capture for the samples was performed using the Roche Nimblegen SeqCap EZ (UW) or Agilent SureSelect Human All Exon 50 Mb (Broad) according the manufacturer's instructions. Paired-end sequencing (2 x 76 bp) was performed using Illumina GAII and HiSeq sequencing instruments. For quality control purposes prior to release of sequence data, samples were initially converted from real-time base-calls to qseq.txt files using Bustard and aligned to hg19 human reference using Burrows-Wheeler Aligner (BWA).[2] Duplicate removal and indel realignment were performed using the Genome Analysis ToolKit (GATK). After using GATK filters, samples were required to reach at least 20x over 70% of the exome target. Prior to release of individual-level sequence reads sequence data were required to match known fingerprint genotypes for their respective samples. Variant calls were evaluated on both bulk and per-sample properties for novel and known variant counts, Ti/Tv ratio, Het/Hom ratio, and Insertion/Deletion ratio. Both bulk and sample metrics were compared to historical values for exome sequencing projects at the two centers. DNA samples that failed laboratory QC were re-queued for library preparation and sequencing.

Variant Calling

Single Nucleotide Variants (SNVs) were called using the UMAKE pipeline at University of Michigan, which allowed all samples to be analyzed simultaneously, both for variant calling and filtering. Briefly, we used BAM[3] files summarizing BWA alignments generated at UW and the Broad as input. These BAM files summarized alignments generated by BWA, refined by duplicate removal, recalibration, and indel re-alignment. We excluded all reads that were not confidently mapped (Phred-scaled mapping quality < 20) from further analysis. To avoid PCR artifacts, we clipped overlapping ends in paired reads. We then computed genotype likelihoods for exome-targeted regions and 50 flanking bases, accounting for per base alignment quality (BAQ) using samtools.[3] Variable sites and their allele frequencies were identified using a maximum-likelihood model, implemented in glfMultiples.[4] These analyses assumed a uniform prior probability of polymorphism at each site. The final call-set was performed on 6,823 samples, referred to as the ESP6800 call-set.

Quality Control

We used a support vector machine (SVM) classifier to separate likely true-positive and false-positive variant sites using a battery of SNP quality metrics. These include allelic balance (the proportional representation of each allele in likely heterozygotes), base quality distribution for sites supporting the reference and alternate alleles, and the distribution of supporting evidence between strands and sequencing cycle, amongst others. We used as the positive training set variants identified by dbSNP or 1000 Genomes, and we used variants that failed multiple filters as the negative training set. We found this method to be effective at removing sequencing artifacts while preserving good-quality data, as indicated by the Ts/Tv ratio for previously known and newly identified variant sites, the proportion of high frequency variants overlapping with dbSNP, and the ratio of synonymous to non-synonymous variants, as well as attempts at validation of a subset of sites. A total of 1,908,614 SNVs passed the SVM filter. To retain only the high quality genotypes, genotypes with a corresponding read depth less than 10 were replaced with a missing value. After doing so, all samples had high call-rates, with the exception of one outlier with a low call-rate. This sample was removed from the analysis. For variants with more than 1 alternate allele, we set to missing any genotype containing a copy of the least frequent alternate allele. Variants with call-rates less than 95% were also removed.

Principal Components Analysis and Ancestry Designation

We conducted a principal component analysis (PCA) to determine sample-level outliers and to cross-check our self-reported ancestry. To do so we only included SNPs with a minor allele frequency (MAF) greater than or equal to 0.1% and a call-rate of greater than 95%. Only autosomal SNPs were included in the PCA. We ran the PCA in PLINK[5] after pruning out SNPs in linkage disequilibrium (LD). The resulting SNPs were used to determine a matrix of genome-wide Identity by State (IBS) pairwise distances, which were subsequently input to the PLINK multidimensional scaling (MDS) algorithm. eFigure 1 shows the first two dimensions from the MDS (analogous to the first two principal components). The first two PCs clearly separate the African American (AA) samples from the European American (EA) samples. However, there is a group of admixed individuals between these two clusters where many self-reported Hispanic individuals were clustered. We removed from all subsequent analyses those individuals of indeterminate genetic ancestry located between the two vertical lines in eFigure 1. For simplicity, we also removed from analysis any individual self-reporting race different from AA or EA. Of the remaining samples, all points to the left of the left-most vertical bar were designated as having AA genetic ancestry. All points to the right of the right-most vertical bar were designated as having EA genetic ancestry. Those samples with discrepant self-reported and designated ancestry were removed from all subsequent analyses.

For each sample we calculated inbreeding coefficients in PLINK. We used the same set of variants that were included in the PCA. One EA sample was found it have an exceedingly high inbreeding coefficient compared to the other samples. This sample was removed from analysis.

Analysis of Relatedness

After designating samples to AA and EA ancestry groups, we ran a race stratified kinship analysis to identify any cryptically related individuals in the ESP6800 call-set. The degree of relatedness was estimated using the KING software.[6] As with the MDS analysis, only LD-pruned autosomal variants with MAF > 0.001 were used as input. Pairs of samples with kinship coefficients consistent with monozygotic twins, first-, second-, and third-degree relatives were excluded when appropriate.

Sex Check

To guard against potential sample swaps, we checked self-reported sex against a normalized measure of read depth on the X and Y chromosomes. Samples where the self-reported sex was clearly different from the XY coverage were considered sample swaps and excluded from further analysis.

GWAS Concordance

When we had access to genome-wide SNP array data, we performed concordance checks between the ESP variants that overlapped with the variants typed on the arrays. Samples identified as having very low concordance rates (<90%) were subsequently dropped from further analysis due to the strong likelihood that they were sample swaps.

eAppendix 3. ExomeChip Genotyping and Quality Control

DNA samples from the WHI clinical coordinating center were sent to the Broad Institute for genotyping and were placed on 96-well plates for processing using the Illumina HumanExome v1.0 SNP array. Genotypes were assigned using GenomeStudio v2010.3. Quality control was performed using the PLINK and R[7] computing platforms. We excluded markers with a genotyping success rate less than 99%. We excluded samples with a genotyping success rate less than 98%. With the resulting sample set, we performed a principal component (PC) analysis as well as an analysis of relatedness using the PLINK IBS/IBD functionality. Outlier samples on the PC plots were excluded, as well as samples thought to be contaminated, based upon results from the relatedness analysis (*i.e.*, they were apparently related to hundreds of other samples). For each related/duplicate pair of samples, we excluded the sample with the lower call-rate. Unexpected duplicate samples were also filtered to prevent potential samples swaps from entering the analysis. For intentionally duplicated samples, we removed samples with low relatedness estimates as we expect them to be close to 1. On the final set of high quality samples, we only included markers with a Hardy-Weinberg Equilibrium $p > 5x10^{-6}$ and call-rate > 90%.

eAppendix 4. Statistical Analyses

The analyses of the ischemic stroke and subtypes involved a total of 2,037 cases and 5,318 controls with exome sequencing or ExomeChip genotype data. A total of 225,239 variants were common to both the exome sequencing and ExomeChip genotyping, of which 119,963 were polymorphic in the full sample. We utilized a single-variant and two gene-based approaches to capture the effects of individual SNPs and the cumulative effects of rare variants (MAF < 5%). For all analyses, age and sex (where appropriate) were included as covariates.

We performed single-variant analyses separately between the sequence and ExomeChip data. For every variant with a MAF ≥ 0.5% (and at least 100 observations with a non-missing genotype), we tested for association between genotype and disease risk with a logistic regression model as implemented in the seqMeta (http://cran.r-prouect.org/web/packages/seqMeta/) package in R. Sequence and ExomeChip specific results were then meta-analyzed with seqMeta software.

We performed gene-level analyses separately between the sequence and ExomeChip data as well. For every missense, nonsense, or splice-variant with MAF <1%, we tested for association with the CMC/T1[8] burden test as implemented in seqMeta. For every missense, nonsense, or splice-variant with MAF <5%, we also tested for association with the SKAT[9] variance components test as implemented in seqMeta. Again, sequence and ExomeChip specific results were then meta-analyzed with seqMeta. Only genes with cumulative MAF > 0.5% were considered in the results.

These analyses were performed using exome sequence data for (a) all ischemic stroke, (b) large vessel subtype, and (c) small vessel subtype. The same models were used for meta-analyses with ExomeChip data, with the same categories but including cardioembolic stroke subtype.

Affected sib-pairs contributed to ESP by the SWISS study were analyzed separately to maximize the impact of the study design. In brief, regions of excess identity by descent (IBD) were examined within and across sib-pairs to identify candidate loci that may be more likely to harbor rare risk variant(s). Annotated non-reference protein coding variants were extracted from the called variants for 237 affected sib-pairs passing QC (529 individuals), focusing on missense, nonsense, stop/gain and stop/loss of function classes. Using code from IBD2.R,[10] single-variant and aggregate IBD probabilities of each gene across all available sib-pairs were estimated and used to generate LOD scores denoting an excess of IBD beyond prior expectations, adjusting for local recombination rates based on European ancestry sample estimates from HapMap Phase3. This estimation of excess IBD at both the variant and gene level was then repeated for (a) 29 small vessel sib-pairs, (b) 18 large artery sib-pairs and (c) 6 cardioembolic sib-pairs. Only TOAST criteria-concordant sib-pairs were included in the subset analyses. Minimal inflation of LOD scores was evident in most analyses, except for the cardioembolic subtypes (eFigure 2). Cardioembolic stroke-related LOD scores were 18.8%-20.8% inflated for the gene-based and single-variant analyses likely due to the few sib-pairs included in these analyses. All other test subsets were conservatively distributed compared to the null distribution of LOD scores (lambda inflation < 1). Based on the results from these analyses, summary statistics from the larger ESP stroke meta-analyses (excluding proband members of affected sib-pairs) were extracted for all gene level and single-variant tests per subtype if a variant or gene had reached a LOD score> 6. P-values for each subset of data were then FDR adjusted, with a reduced penalty for multiple testing based on a minimum LOD of 6 in the affected sib-pair analysis. Only 1 variant in *PDE4DIP* passed correction for multiple testing.

References

1. Tennessen JA, Bigham AW, O'Connor TD, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*. 2012;337(6090);64-69.

2. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595.

3. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.

4. Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008;18(11):1851-1858.

5. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet*. 2007;81(3):559-575.

6. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867-2873.

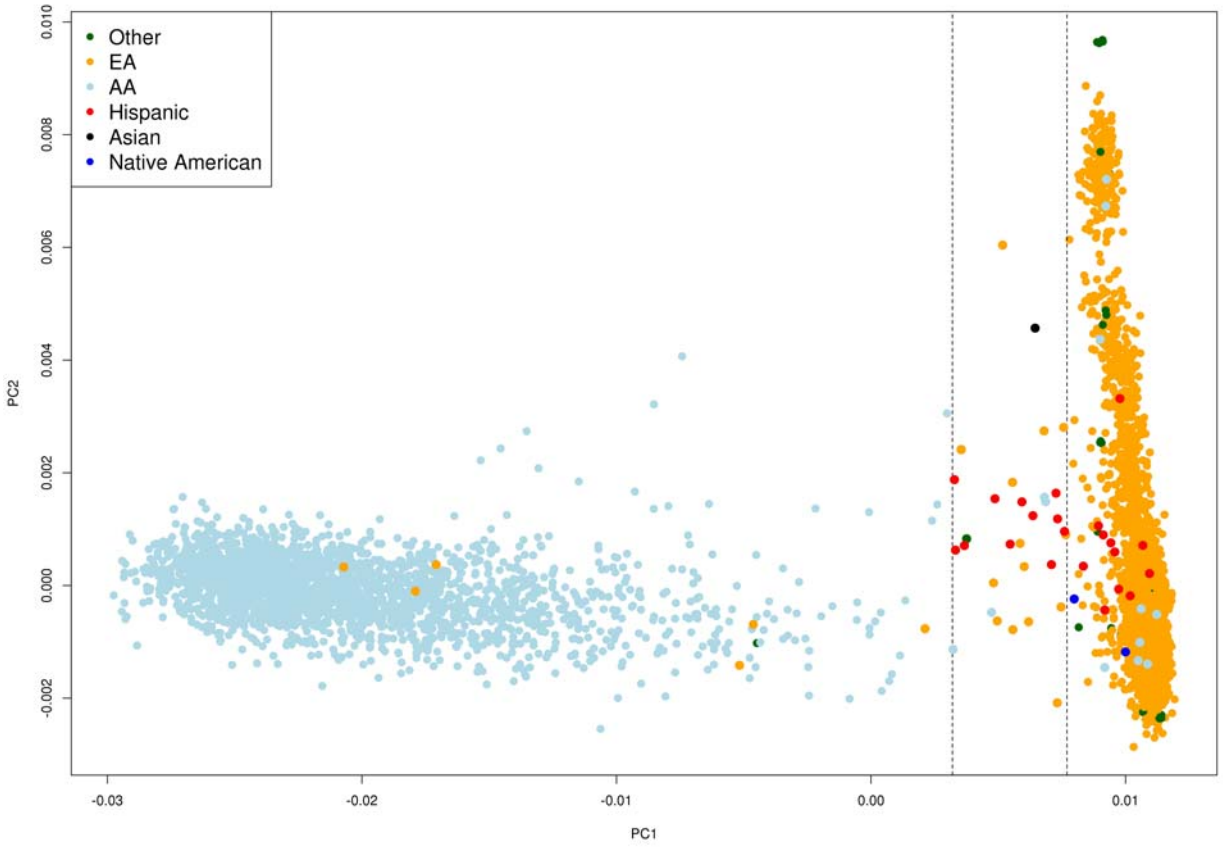
7. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.ISBN 3-900051-07-0, URL http://www.R-project.org/. 2012.

8. Li B and Leal SM Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet*. 2008;83(3):311-321.

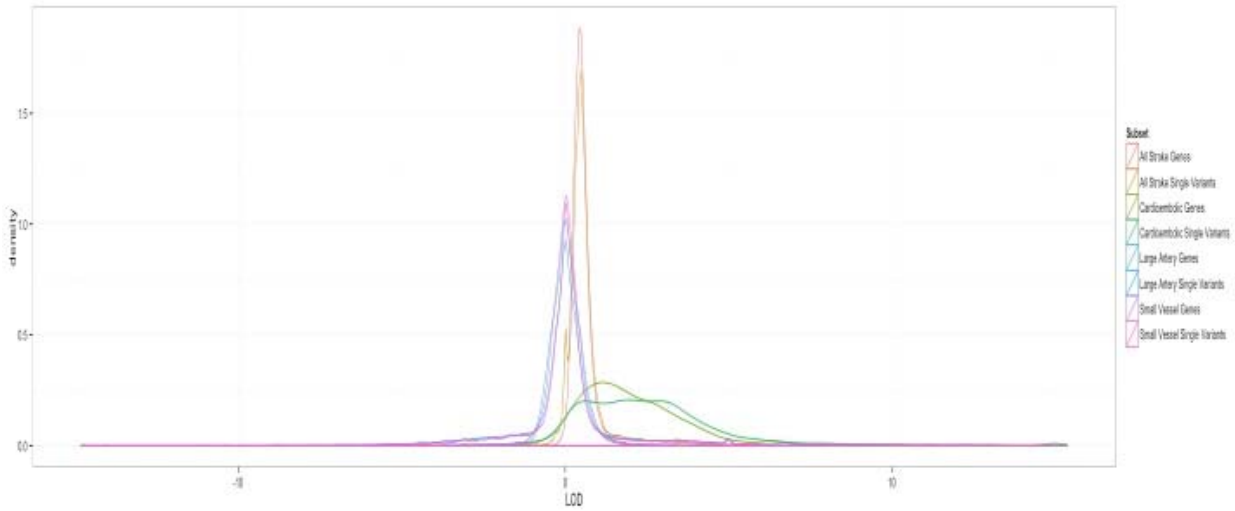
9. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. *Am J Hum Genet*. 2011;89(1):82-93.

10. Krawitz PM, Schweiger MR, Roedelsperger C, et al. Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nat Genet*. 2010;42(10):827-829.

eFigure 1. Principal Components 1 (PC1) and 2 (PC2) from the ESP6800 with self-reported ancestry of European Caucasian (EA, orange), African (AA, light blue), Hispanic (red), Asian (black), and Native American (dark blue); and missing on self report (Other, green).



eFigure 2. Density distribution of LOD scores per subset of IBD tests in affected sib pairs with ischemic stroke.



eTable 1. Meta-analysis of single-variant and gene-based association results with ischemic stroke for previously published candidate genes/loci.

		Single Variant Test				Gene-based test (SKAT)		
Gene	Location	SNV	MAF	OR	p_m	Number of SNVs	Cumulative MAF	p-value
MTHFR	1p36.3	rs1801133	0.351	0.938	0.507	9	0.077	0.488
F5	1q23	rs6032	0.277	1.381	0.002	18	0.069	0.299
PMVK	1q21.3	rs139248801	0.007	1.498	0.065	7	0.010	0.071
ABL2	1q25.2	rs1318056	0.083	0.932	0.301	8	0.020	0.691
PPARG	3p25	rs1801282	0.120	0.912	0.109	1		
F13A1	6p25.3	rs5985	0.259	0.991	0.832	7	0.038	0.523
LPA	6q26	rs41267809	0.024	0.830	0.122	25	0.128	0.290
PON1	7q21.3	rs854560	0.372	1.040	0.312	8	0.007	0.309
PIK3CG	7q22.3	rs61749915	0.018	0.724	0.020	9	0.069	0.022
NOS3	7q36					14	0.011	0.016
LPL	8p21.3	rs301	0.261	0.911	0.033	6	0.036	0.559
PINX1	8p23	rs1078543	0.134	0.733	0.027	7	0.023	0.822
ZHX2	8q24.13	rs142123946	0.005	0.693	0.149	8	0.080	0.347
CDKN2A	9p21.3	rs3731248	0.030	0.972	0.797	3	0.033	0.832
CDKN2B-AS1	9p21.3	rs564398	0.401	0.959	0.320	0		
C9orf3	9q22.32					15	0.005	0.039
ABCA1	9q31.1	Rs2066715	0.060	1.340	$2x10^{-4}$	25	0.051	0.489
TLR4	9q33.1	rs4986790	0.060	0.895	0.156	10	0.010	0.416
SYNPO2L	10q22.2	rs60632610	0.148	0.903	0.076	4		
PDCD11	10q24.33	rs61751511	0.014	0.703	0.028	47	0.064	0.038
APOA5	11q23.3	rs3135506	0.059	1.133	0.118	7	0.027	0.237

eTable 1. Meta-analysis single-variant and gene-based association results with ischemic stroke for previously published candidate genes/loci. (continued)

		Single Variant Test				Gene-based test (SKAT)		
Gene	Location	SNV	MAF	OR	p_m	Number of SNVs	Cumulative MAF	p-value
LTA4H	12q22	rs143821623	0.007	1.106	0.659	5	0.008	0.652
SCARB1	12q24.31	rs5891	0.010	0.877	0.477	7	0.014	0.624
PRKCH	14q23.1	rs2230500	0.009	0.616	0.316	5	0.011	0.520
SYNE2	14q23.2	rs17751301	0.072	1.151	0.051	98	0.294	0.432
HCN4	15q24.1					6	0.007	0.773
ZFHX3	16q22.3	rs62639999	0.052	0.867	0.090	49	0.126	0.061
GP1BA	17p13.2	rs6065	0.083	0.980	0.768	0		
ITGB3	17q21.32	rs5918	0.154	1.017	0.740	0		
FBF1	17q25.1	rs113062332	0.029	1.223	0.068	11	0.039	0.134
ACOX1	17q25.1	rs1135640	0.352	1.054	0.177	9	0.005	0.508
TRIM65	17q25.1	rs61754864	0.012	1.257	0.228	4	0.018	0.342
MRPL38	17q25.3	rs9191	0.017	0.908	0.498	8	0.041	0.564
NOTCH3	19p13.12	rs112197217	0.017	0.645	0.212	15	0.039	0.771
APOE	19q13.2	rs7412	0.083	0.969	0.673	0		
MACROD2	20p12.1	rs2990505	0.218	1.042	0.368	6	0.006	0.458
FLRT3	20p12.1	rs35253731	0.025	0.811	0.084	4	0.061	0.258
CBS	21q22.3	rs117687681	0.005	0.790	0.398	13	0.013	0.483
GLA	Xq22.1					5	0.007	0.608

** Genes interrogated by exome sequencing but without SNVs having MAF > 0.005 or total (cumulative) variant MAF > 0.005

eTable 2. Meta-analysis of single-variant and gene-based association results with ischemic stroke and small vessel subtype for previously published candidate genes/loci.

		Single Variant Test				Gene-based test (SKAT)		
Gene	Location	SNV	MAF	OR	p_m	Number of SNVs	Cumulative MAF	p-value
C9orf3	9q22.32					13	0.005	0.626
SYNPO2L	10q22.2	rs34163229	0.160	0.977	0.877	0		
PDCD11	10q24.33	rs61751511	0.014	0.516	0.033	35	0.066	0.100
SYNE2	14q23.2	rs117070973	0.006	2.140	0.121	98	0.296	0.533
HCN4	15q24.1					6	0.006	0.378
TRIM65	17q25.1	rs61754864	0.011	0.979	0.965	4	0.017	0.855
FBF1	17q25.1	rs2305913	0.348	0.911	0.218	11	0.038	0.039
ACOX1	17q25.1	rs1135640	0.347	0.926	0.308	0		
MRPL38	17q25.3	Rs34136221	0.018	1.305	0.471	8	0.042	0.961
NOTCH3	19p13.12	rs112197217	0.009	1.818	0.101	13	0.039	0.106
MACROD2	20p12.1	rs2990505	0.216	1.010	0.907	0		
FLRT3	20p12.1	rs8120693	0.008	0.519	0.113	4	0.063	0.320
GLA	Xq22.1					5	0.008	0.887

eTable 3. Meta-analysis of single-variant and gene-based association results with ischemic stroke and large vessel subtype for previously published candidate genes/loci.

		Single Variant Test				Gene-based test (SKAT)		
Gene	Location	SNV	MAF	OR	p_m	Number of SNVs	Cumulative MAF	p-value
MTHFR	1p36.3	rs1801133	0.344	0.899	0.317	9	0.076	0.753
ABL2	1q25.2	rs17277288	0.014	0.758	0.504	8	0.021	0.007
PPARG	3p24	rs1801282	0.122	0.791	0.125	0		
LPA	6q26	rs41272110	0.155	0.796	0.2533	25	0.131	0.885
PIK3CG	7q22.3	rs61749915	0.046	2.076	0.027	9	0.069	0.201
NOS3	7q36					14	0.011	0.348
PINX1	8p23	rs1078543	0.141	0.751	0.273	7	0.023	0.848
ZHX2	8q24.13	rs142123946	0.031	0.57	0.048	7	0.080	0.135
CDKN2A	9p21.3	rs3731249	0.031	1.485	0.186	3	0.034	0.222
CDKN2B-AS1	9p21.3	rs564398	0.502	0.753	0.046	0		
TLR4	9q33.1	rs4985690	0.061	0.860	0.478	10	0.011	0.585
LTA4H	12q22	rs143721623	0.007	2.526	0.146	4	0.007	0.164
SCARB1	12q24.31	rs5891	0.010	0.659	0.390	6	0.014	0.621
PRKCH	14q23.1	rs2230500	0.009	0.496	0.330	5	0.011	0.400
APOE	19q13.2	rs7312	0.084	1.306	0.289	0		
CBS	21q22.3	Rs117687681	0.006	4.048	0.121	13	0.013	0.255

eTable 4. Meta-analysis of single-variant and gene-based association results with ischemic stroke and cardioembolic subtype for previously published candidate genes/loci.

		Single Variant Test				Gene-based test (SKAT)		
Gene	Location	SNV	MAF	OR	p_m	Number of SNVs	Cumulative MAF	p-value
PMVK	1q21.3	rs139248801	0.006	1.662	0.193	5	0.007	0.214
ZFHX3	16q22.3	rs149133285	0.005	0.324	0.010	40	0.102	0.411

Gene Names in eTable1-eTable4

MTHFR (methylenetretrahydrofolate reductase (NAD(P)H), 1p36.3); *F5* (coagulation factor V (proaccelerin, labile factor), 1q23); *PMVK* (phosphomevalonate kinase, 1q21.3); *ABL2* (v-abl Abelson murine leukemia viral oncogene homolog 2, 1q25.2); *PPARG* (peroxisome proliferator-activated receptor gamma, 3p25); *F13A1* (coagulation factor XIII, A1 polypeptide, 6p25.3); *LPA* (lipoprotein, Lp(a), 6q26); *PON1* (paraoxonase 1, 7q21.3); *PIK3CG* (phosphoinositide-3-kinase, catalytic, gamma polypeptide, 7q22.3); *NOS3* (nitric oxide synthase 3 (endothelial cell), 7q36); *LPL* (lipoprotein lipase, 8p21.3); *PINX1* (PIN2/TERF1 interacting, telomerase inhibitor 1, 8p23); *ZHX2* (zinc fingers and homeoboxes 2, 8q24.13); *CDKN2A* (cyclin-dependent kinase inhibitor 2A, 9p21.3); *CDKN2B-AS1* (CDKN2B antisense RNA 1, 9p21.3); *C9orf3* (chromosome 9 open reading frame 3, 9q22.32); *ABCA1* (ATP-binding cassette, sub-family A (ABC1), member 1, 9q31.1); *TLR4* (toll-like receptor 4, 9q33.1); *SYNPO2L* (synaptopodin 2-like, 10q22.2); *PDCD11* (programmed cell death 11, 10q24.33); *APOA5* (apolipoprotein A-V, 11q23.3); *LTA4H* (leukotriene A4 hydrolase, 12q22); *SCARB1* (scavenger receptor class B, member 1, 12q24.31); *PRKCH* (protein kinase C, eta, 14q23.1); *SYNE2* (spectrin repeat containing, nuclear envelope 2, 14q23.2); *HCN4* (hyperpolarization activated cyclic nucleotide-gated potassim channel 4, 15q24.1); *ZFHX3* (zinc finger homeobox 3, 16q22.3); *GP1BA* (glycoprotein 1b (platelet), alpha polypeptide, 17p13.2); *ITGB3* (integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61), 17q21.32); *FBF1* (Fas (TNFRSF6) binding factor 1, 17q25.1); *ACOX1* (acyl-CoA oxidase 1, palmitoyl, 17q25.1); *TRIM65* (tripartite motif containing 65, 17q25.1); *MRPL38* (mitochondrial ribosomal protein L38, 17q25.3); *NOTCH3* (notch 3, 19p13.12); *APOE* (apolipoprotein E, 19q13.2); *MACROD2* (MACRO domain containing 2, 20p12.1); *FLRT3* (fibronectin leucine rich transmembrane protein 3, 20p12/1); *CBS* (cystathionine-beta-synthase, 21q22.3); *GLA* (galactosidase, alpha, Xq22.1)

Genes (and their location) not having variants in coding regions by exome sequencing or on the ExomeChip (large vessel stroke, subclinical atherosclerosis, small vessel stroke, atrial fibrillation, white matter disease, or stroke risk factors):

KCNN3 (potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3, 1q21.3); *PRRX1* (paired related homeobox 1, 1q24.2); *TREX1* (three prime repair exonuclease 1, 3p21.31); *PITX2* (paired-like homeodomain 2, 4q25); *EDNRA* (endothelin receptor type A, 4q31.22); *PDE4D* (phosphodiesterase 4D, cAMP-specific, 5q12.1); *WNT8A* (wingless-type MMTV integration site family, member 8A, 5q31); *HDAC9* (histone deacetylase 9, 7p21.1); *CAV1* (caveolin 1, caeolae protein, 22kDa, 7q31.1); *ACTA2* (actin, alpha 2, smooth muscle, aorta, 10q23.3); *HTRA1* (HtrA serine protease 1, 10q26.3); *HBB* (hemoglobin, beta, 11p15.5); *NINJ2* (ninjurin 2, 12p13.33); *ALOX5AP* (arachidonate 5-lipoxygenase-activating protein, 13q12.3); *WBP2* (WW domain binding protein 2, 17q25.1); *TRIM47* (tripartite motif contain 47, 17q25.1); *APOC1* (apolipoprotein C-1, 19q13.2); *APP* (amyloid beta (A4) precursor protein, 21q21.3)

