Exome sequences were obtained for 1,376 FALS cases and 13,883 controls. Samples were excluded in the event of exome-wide call rate<70%, outlying heterozygosity (F<0.1 or F>0.1), SNP predicted and reported gender discrepancy, detectable relatedness to another retained sample (Kinship coefficient>0.0442; ≥3rd degree relationship), outlying ancestry with respect FALS samples in pairwise tests of population concordance (exhibits P<1x10^-4 in tests with ≥10% of FALS cases; Supplementary Fig. 2) or outlying ancestry with respect FALS samples in subsequent principle components analysis (eigenvector value >4 standard deviations from FALS mean along any of principle components 1-4).
Supplementary Figure 2

Stratification analysis of FALS discovery cohort.

a) Results from first round of population outlier filtering. Y-axis denotes the proportion of FALS samples for which a given test sample exhibits significant population discordance ($P<1.0 \times 10^{-4}$ in pairwise population concordance testing). X-axis displays corresponding geographical labels for FALS cases. Horizontal dotted line denotes 10% FALS discordance threshold, all cases and controls falling above this line were removed during first round of stratification filtering. b) Distribution of FALS samples along eigenvectors 1 and 2 following principle components analysis of QC filtered FALS discovery cohort. c) Distribution of cases and controls along eigenvectors 1 and 2 following principle components analysis of QC filtered FALS discovery cohort. AUS Australia; BEL Belgium; CAN Canada; ESP Spain; GER Germany; IRL Ireland; ITA Italy; NLD Netherlands; TUR Turkey; UK United Kingdom; USA United States; USA_AFR African American; USA_AMR Ad mixed American.
To identify loci potentially subject to confounding bias in FALS RVB analyses, RVB were performed across all known potential sources of heterogeneity in the FALS control cohort. This involved dividing controls into 28 distinct pseudo case-control groups based on the sequencing center and associated project to identify loci showing association with non-ALS related data, population or phenotypic stratifiers. Y-axis denotes P-values observed during ALS gene trained RVB testing in FALS vs controls. X-axis denotes minimum P-value observed during ALS gene trained RVB testing in the 28 pseudo case-control cohorts. Genes shown in grey achieve $P<1\times10^{-3}$ for possible confounder association. Known and candidate ALS genes show no confounder association.

Supplementary Figure 3

Control – control analyses.
Supplementary Figure 4

*NEK1* discovery cohort coverage.

Plot of variant call rate across *NEK1* protein coding region in cases vs controls.
Supplementary Figure 5

Inbreeding coefficients from Dutch whole genome sequencing cohort.

Four ALS patients sampled from an isolated community in the Netherlands can be seen to exhibit elevated coefficients of inbreeding (shown in red) relative to a larger panel of Dutch genome sequences (n=1861). Boxplot shows cohort median, interquartile range, 2.5% quantile and 97.5% quantile.
Supplementary Figure 6

Autozygosity mapping reveals NEK1:p.R261H as candidate ALS variant.

Whole genome sequencing followed by autozygosity mapping with allowed genetic heterogeneity identified 10 runs of homozygosity present in one or more of four SALS patients from an isolated Dutch community (Top panel). These regions contained four variants where at least one of the four patients was homozygous and where MAF was less than 0.01 in the 1000 genomes project, the NHLBI exome sequencing project\textsuperscript{52} and ExAC\textsuperscript{53} (bottom panel). NEK1:p.R261H is the only variant identifiable in all patients and the only variant for which multiple homozygous genotypes were observed.
Supplementary Figure 7

QC of NEK1 LOF and p.R261H SALS replication cohorts.

Full NEK1 sequencing was performed for 2,387 SALS and 1,093 matched controls. p.R261H genotypes were obtained for 8,173 SALS and 5,189 controls (inclusive of 2,387 SALS and 1,093 controls with full NEK1 sequencing). Samples excluded in the event of outlying heterozygosity (F<0.1 or F>0.1), SNP predicted and reported gender discrepancy, detectable relatedness to a sample from the FALS cohort or retained sample from SALS replication cohort (Kinship coefficient>0.0884, ≤2nd degree relationship), outlying ancestry as assessed by identity by state distance to 5th nearest neighbour (>3 standard deviations from group mean) or outlying ancestry as assessed by principle components analysis (eigenvector value >4 standard deviations from group mean along any of principle components 1-4).
Supplementary Figure 8

Stratification analysis of SALS replication cohorts.

a-b) Distribution of cases and controls along eigenvectors 1 and 2 following principle components analysis of QC filtered NEK1 LOF replication cohort. c-d) Distribution of cases and controls along eigenvectors 1 and 2 following principle components analysis of QC filtered NEK1 p.R261H replication cohort. BEL Belgium; ESP Spain; GER Germany; IRL Ireland; ITA Italy; NLD Netherlands; UK United Kingdom; USA United States.
Supplementary Figure 9

Distribution of NEK1 variants.

Observed case-control distribution of NEK1 variants in FALS (a) and SALS (b) cohorts. LOF variants are highlighted in black, missense variants are labeled in grey. HGVS descriptions are followed by case/control carrier counts in parenthesis. Predicted splice altering variants are indicated with an asterisk.