nature research

Corresponding author(s): Kelly A. Frazer

Last updated by author(s): Jul 8, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Software was not used for data collection
Data analysis	The following software were used: BWA-MEM (0.7.15); GATK(3.4-46); STAR (2.5.0a); RSEM(1.2.20); MACS2 (2.1.0); WASP (0.2.2); HOMER (4.7); fgwas; PLINK(1.07); GREGOR(v1.4.0); HICCUPS (1.6), Fit-Hi-C (1.0.1), DESeq2 (R 3.2.2), featureCounts (1.5.0), MEME-CHIP (online version), bedtools (2.25.0), Python (2.7.11), CHOPCHOP (online version, 2.0.0), Nexus Copy Number software (7.5). All statistical analyses were performed using R, version 3.2.2. Custom codes and Jupyter notebooks are available at https://github.com/frazer-lab/NKX2-5_ASE_iPSC-CM

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All iPSC lines are available through WiCell Research Institute (www.wicell.org; NHLBI Next Gen Collection). All genomic data are available through dbGAP accessions phs000924 (RNA-Seq, ChIP-Seq, ATAC-Seq, Hi-C) and phs001325 (whole-genome sequence SNV and CNV genotypes), NCBI BioProject PRJNA285375. Processed data files are available through GEO accessions GSE125540 and GSEXXXXXX.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample sizes. We selected seven individuals of Asian and European descent in the iPSCORE resource that are part of a three-generational family. Our study design included three genetically unrelated subjects and two parent-offspring quartets, which enabled us to examine the inheritance of genetic effects. The primary analyses in this manuscript identify differences in molecular phenotypes between two alleles within an individual (ASE – allele-specific effects). Therefore, the power for these analyses was primarily dependent on the read depth of RNA-Seq, ChIP-Seq or ATAC-Seq at heterozygous SNVs, rather than the number of subjects. To improve power we then combined the results from the same SNV across individuals using a meta-analysis.
Data exclusions	We used established quality control metrics and filtering criteria for all molecular data: For all sequencing data, blacklisted regions from ENCODE, reads mapping in chromosome other than chr1-chr22, chrX, chrY, and read-pairs with mapping quality Q<30 were filtered. For ChIP-Seq and ATAC-Seq, peaks were filtered for q-value <0.01, and samples with a FrIP <4% were excluded.
Replication	We compared the ASE (allele-specific effects) between the different molecular data, which showed good correlation, indicating that the genetic effects are reproducible. Using our approach we identified at least one variant (at the SCN10A locus) that was previously described to be functional and showing allele-specific binding. EMSA experiment for variants at the CAV1 and GNB4 loci were performed twice and showed consistent results in both independent experiments, as shown in the paper's figures; for the variant at the SSBP3 locus, we performed 4 EMSA experiments, of which 3 showed consistent allelic differences (two of those are shown in the paper) and one showed no difference. CRISPRi was performed in two independent cell lines and the results were reproducible. Luciferase assays had 6 replicates each variant, consisting of transfection of 6 different wells.
Randomization	As we are testing for genetic associations, there are no experimental groups in this study
Blinding	As there are no experimental groups in this study, blinding was not applicable for this study

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-TNNT2 (Thermo Scientific MA5-12960, clone 13-11, dilution 1:200, monoclonal) anti -MYL7 (Synaptics Systems 311011, clone 56F5, dilution 1:200, monoclonal) anti-H3K27ac (Abcam ab4729, lots GR183922-2 (1.75µg per IP), GR184333-2 (1µg per IP), or GR00324078 (1µg per IP), polyclonal) anti- NKX2-5 (Santa Cruz Biotechnology, sc-8697x, lot C0113, 5µg per IP, polyclonal)
Validation	anti-TNNT2: from company website: MA5-12960 targets Troponin T Cardiac Isoform in IF/ICC, IHC (P), and IM applications and shows reactivity with Avian, Canine, Chicken, Fish, Guinea Pig, Human, mouse, Porcine, Rabbit, and Rat samples. Referenced in 122 publications. anti -MYL7:from company website: Tested applications: WB, ICC, IHC, IHC-P/FFPE, Reacts with: human (Q01449), rat, mouse

(Q9QVP4). No signal: chicken. Referenced in 36 publications.
anti-H3K27ac: from company website: Suitable for: IHC-Fr, ICC/IF, WB, IHC-P, CHIPseq, ChIP/Chip, ChIP, PepArr. Reacts with: Mouse, Rat, Chicken, Cow, Human, Arabidopsis thaliana, Drosophila melanogaster, Monkey, Zebrafish, Plasmodium falciparum, Rice, Cyanidioschyzon merolae. Referenced in 766 publications.
anti- NKX2-5: from company website: recommended for detection of Nkx-2.5 and, to a lesser extent, Nkx-2.3 of mouse, rat and human origin by WB, IP, IF and ELISA; also reactive with additional species, including and canine, bovine and porcine. Referenced in 30 publications. Nkx-2.5 (N-19) has been discontinued and replaced by Nkx-2.5 (A-3): sc-376565

Additionally, anti-NKX2-5 (lot C0113) was further validated in our laboratory by ChIP followed by Western Blot in iPSC-CMs.

Human research participants

Policy information about studies involving human research participants

Population characteristics	All subject information is provided in Supplementary Table 1. In summary, a family of seven individuals spanning three generations was utilized, with all individuals from EAS, EUR, or EAS/EUR descent, spanning ages 18-77. Five members of this family segregate the long-QT syndrome type 2 mutation KCNH2 p.Trp1001* (rs121912509, c.3003G>A), and the genotypes are reported. However, the disease phenotype is not analyzed in this study. These individuals are explained in Panopolous et. al, as cited in the references. Individual covariates (sex, ethnicity, disease state) were not used in the analysis of NKX2-5 allele-specific effect.
Recruitment	Recruitment for these individuals is explained fully in Panopolous et. al, as cited in the references. Specifically, these 7 individuals were recruited through both the Twin Sibling Pedigree cohort (TSP; a population-based twin registry spanning counties in Southern California) and open enrollment through the Clinical and Translational Research Institute (CTRI) at the University of California at San Diego (UCSD). Each of the subjects first consented to the study and filled out a questionnaire. These data were transcribed to a database and subjects were de-identified with a new sample ID. Ethnicity was reported as a free-response answer and translated into one of six recorded ethnicity groupings (African American, European, Hispanic, Indian, Middle Eastern, Asian). A seventh category was used when more than one ethnicity was reported; that individual was recorded as "Multiple ethnicities reported." We do not report any recruitment or self-selection bias that could have influenced the results of this study.
Ethics oversight	This study was approved by the Institutional Review Boards of the University of California at San Diego (Project #110776ZF).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	BAM files for all ChIP-seq data are available through dbGAP (phs000924, BioProject PRJNA285375), as it contains identifiable information. BigWig files for H3K27ac were deposited in GEO (GSE125540), as a part of a parallel publication utilizing the same data (Greenwald, W.W. and Li, H. et al. Subtle changes in chromatin loop contact propensity are associated with differential gene regulation and expression. Nat Commun, 2019). The NKX2-5 ChIP-Seq BigWig files were deposited in GEO, accession GSE133833
Files in database submission	NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_1_FS009 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_3_FS010 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_3_FS003 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_6_FS018_A NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_2_FS005_A NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_4_FS015_A NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_9_FS008 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_1_FS016 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_7_FS011 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_4_FS015_B NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_3_FS024 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_9_FS007 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_6_FS018_B NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_7_FS014 NKX2-5_ChIPSeq_iPSC-CM_iPSCORE_2_2_FS005_B NKX2-5_ChIPSeq_iPSC-CM_Pool_Input control 036253ba-25d4-4b93-9127-adc21527a082_iPSC-CM (ChIP-seq) 03a95c32-f31e-4846-b4e7-6991a7bbbd86_iPSC (ChIP-seq) 0441cd83-9d1a-416b-a82d-856f7f04f6e4_iPSC-CM (ChIP-seq) 079927aa-29e1-462c-8b15-af1a3d8f1b20_iPSC (ChIP-seq) 093b028a-b5e8-496e-b2cb-aa5e0d4da368_iPSC (ChIP-seq) 190c0665-f7ff-48af-925f-a50607fe9af3_iPSC (ChIP-seq) 1a75c25e-c463-47ab-b838-0fdf7af5ea24_iPSC-CM (ChIP-seq) 20f37c64-0c86-4cb7-9501-6132d1801b84_iPSC-CM (ChIP-seq)

232f80c9-44f3-45e5-ac7c-88c3b27950ea_iPSC-CM (ChIP-seq)
262ead42-62d0-45fb-a309-dfb9fdeceb28_iPSC-CM (ChIP-seq)
2d5dda6b-f7cb-428b-8c85-380325c4926a_iPSC (ChIP-seq)
31ae89ae-0296-415c-b202-e9668ab4461c_iPSC (ChIP-seq)
32ddc48c-0759-4444-ad25-465c591b9d7a_iPSC (ChIP-seq)
36f8c31d-b865-4424-a829-8e55cbba6411_iPSC (ChIP-seq)
4a049ce8-ac3f-4d2b-8bc8-76a2c376bed9_iPSC (ChIP-seq)
4ee11143-6657-4375-b56e-51b1c35c8f3d_iPSC (ChIP-seq)
6a65308f-37d0-4b06-94b3-c5301d760afd_iPSC-CM (ChIP-seq)
6d634ac7-1854-4d2c-b7db-1d9913dc8dc6_iPSC-CM (ChIP-seq)
76a804b4-b3ec-4b54-9c70-92b11cf82f33_iPSC (ChIP-seq)
7a1966dd-3453-4d30-8e77-6870ee9cd790_iPSC-CM (ChIP-seq)
7ba84fca-758f-4068-a7ef-914e68be9c3e_iPSC-CM (ChIP-seq)
7d2f5a9b-831d-4859-9451-3c7dcc0a00ed_iPSC-CM (ChIP-seq)
84943ee6-4aef-468e-ba30-55d150e879b5_iPSC-CM (ChIP-seq)
881424e7-e3cb-482c-830b-69c6897eb772_iPSC-CM (ChIP-seq)
955ebdbe-26c8-43c3-8a1d-207728297dc0_iPSC (ChIP-seq)
960c428e-6eba-4fd6-b926-f8776fc20cbc_iPSC (ChIP-seq)
962cd048-2790-4ef4-8c06-acf9ec1b7bc2_iPSC-CM (ChIP-seq)
a1d6f499-3f5f-494f-a182-1e3211bd5ae1_iPSC-CM (ChIP-seq)
a31d0ae6-a7ef-4054-98f0-2928ccf16cf4_iPSC-CM (ChIP-seq)
a80bc6f4-b918-4461-ab7b-c1f16136bed6_iPSC-CM (ChIP-seq)
a8241e50-58c1-40de-9551-573d41ea4f19_iPSC (ChIP-seq)
aa613f3f-2ff5-4d09-84bc-008c30c6ef66_iPSC (ChIP-seq)
ac83d79a-3621-4eb2-8245-5bcccc209d58_iPSC-CM (ChIP-seq)
aeeaf78e-a3ae-4071-b364-5d4e35e06799_iPSC (ChIP-seq)
b389c69b-47cc-4b1e-ae53-bc0a8b23f88a_iPSC-CM (ChIP-seq)
b56fb523-1e03-4812-a745-1f97314359e7_iPSC-CM (ChIP-seq)
c92b23f5-206f-4892-ba1b-90dfb8cfe2ee_iPSC-CM (ChIP-seq)
cc566b8a-fc54-4941-8328-f57401635839_iPSC-CM (ChIP-seq)
d08fcfb6-1540-4174-bc73-b625da9d9ab9_iPSC-CM (ChIP-seq)
d56606a1-e263-4da9-acf9-2d6f14a822cb_iPSC-CM (ChIP-seq)
d7a034e7-2916-409d-9197-51c6b3b8e173_iPSC-CM (ChIP-seq)
dd8bcb5e-15bf-4be6-97ac-94f54ed811eb_iPSC (ChIP-seq)
e241faad-584b-41b5-9d1d-07f8ce5ecd30_iPSC (ChIP-seq)
e86d9c4d-d0ca-4122-b729-9b1f0fbee934_iPSC-CM (ChIP-seq)
e8d0daf3-5f47-44eb-afdd-f060c0d8a3f9_iPSC (ChIP-seq)
e9586b34-11fc-4cdf-907a-beb5643ee3ae_iPSC-CM (ChIP-seq)
f0f2cac5-03e9-4334-8dfe-0d4c13c9a511_iPSC (ChIP-seq)
f8db3685-37ed-409c-8216-c0d045108403_iPSC (ChIP-seq)

Genome browser session
(e.g. UCSC)

https://genome.ucsc.edu/s/PaolaB/Benaglio_NKX2%2D5_ChIPSeq_public
https://genome.ucsc.edu/s/PaolaB/Benaglio_H3K27ac_iPSC%2DCMs_public
https://genome.ucsc.edu/s/PaolaB/Benaglio_H3K27ac_iPSCs_public

Methodology

Replicates

We generated and analyzed 48 ChIP-Seq of histone modification H3K27ac (iPSCs: 17 samples and 4 technical replicates; iPSC-CMs: 25 samples and 2 technical replicates), and 15 ChIP-seq of NKX2-5 (iPSC-CMs: 12 samples and 3 technical replicates) as detailed in Supplementary Tables 2 and 3.
The median pairwise Spearman correlation coefficient across all samples and across replicate samples of the same individual were, respectively :
0.97 and 0.98 for H3K27ac in iPSCs, 0.93 and 0.93 for H3K27ac in iPSC-CMs and 0.79 and 0.79 for NKX2-5

Sequencing depth

All details on individual sequencing metrics are provided in Supplementary table 2 and summarized in Supplementary table 3. For NKX2-5 the average depth was 26 M uniquely mapped reads while for H3K27ac was 27 M for iPSC-CMs, and 40 M for iPSCs. We used in 100 PE reads in most of cases, and 150 PE in few cases.

Antibodies

anti-H3K27ac (Abcam ab4729, lots GR183922-2 (1.75µg), GR184333-2 (1µg), or GR00324078 (1µg))
anti- NKX2-5 (Santa Cruz Biotechnology, sc-8697x, lot C0113, 5µg)

Peak calling parameters

Peak calling was performed using MACS2 ('macs2 callpeak -f BAMPE -g hs -B --SPMR --verbose 3 --cutoff-analysis --call-summits -q 0.01') with reads derived from sonicated chromatin not subjected to IP (i.e. input chromatin) from a pool of samples used as negative control. Peak coordinates were called from combined samples of either iPSCs or iPSC-CMs, generated by pooling the BAM files of each data type across all samples of the given cell type.

Data quality

All peaks provided and utilized in the analyses were FDR (q) < 0.01
Motif enrichment analysis was performed using HOMER 'findMotifsGenome.pl' and, for NKX2-5, also using MEME ChIP. Motif analysis of the NKX2-5 ChIP-Seq confirmed a significant enrichment (binomial test, q-value <0.0001) for the NKX2-5 homeobox motif, as well as for the motifs of other heart development TFs (GATA4, TBX5, TBX20, MEF2A/C and MEIS1).

Software

Alignment: BWA
Peak calling: MACS2
Quantification of signal: featureCounts

Motif enrichment: HOMER and MEME
Allele-specific effect analysis: WASP, MBASED, GATK
Postprocessing analyses and data manipulation: R custom codes found at https://github.com/frazer-lab