**Supplementary Table 2**: Prioritization of 1048 test genes located on chromosome 3 using training genes of congenital heart defects (CHD), arrhythmias (AR), and cardiomyopathies (CM)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Known or putative involvement in congenital heart defects</th>
</tr>
</thead>
</table>
| EVI1   | The Evi1 proto-oncogene is required at midgestation for neural, heart, and paraxial mesenchyme development indicating a role in neural crest cells
dual role in heart development |
| ITGB5  | β5-integrin is a key integrin involved in angiogenesis, vasculogenesis, hematopoiesis and also plays a role in the proliferation of cardiac cells
| GATA2  | Belongs to GATA family of TFs (GATA4 is one of the training genes)
| FOXP1  | Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation in mice
| SHOX2  | SHOX genes are involved in Turner syndrome, which is a condition associated with reduced final height and gonadal dysgenesis. A number of other signs and symptoms are seen more frequently with the syndrome. With respect to cardiac function, congenital malformations of the heart and the great vessels
| EOMES  | Transcription factor of the T-box gene family
| FBLN2  | Extracellular matrix protein fibulin-2 is expressed in the embryonic endocardial cushion tissue and is a prominent component of valves in adult heart
| ZIC1   | Mutations in ZIC1 cause Dandy-Walker malformation (DWM; OMIM #220200), which is a common but poorly understood congenital cerebellar malformation. There is an increased frequency with congenital heart disease, cleft lip/palate and neural tube defects
| PLXNA1 | Plays a role in cardiac chamber formation: deficiency of PLXNA1 leads to a thin ventricular layer and to defective trabeculation of the heart
| RARB   | RARB is a receptor for retinoic acid and either a deficit or an excess of retinoic acid may result in congenital birth defects
| SEMA3B | Semaphorin family is implicated in neural crest guidance and SEMA3E is a also trainings gene
| SEMA3E | Semaphorin 3E is a training gene
| HYAL2 and HYAL3 | Hyaluronidases are known to play a role in cardiac cushion migration during the development of the heart
| *      | 16 uncharacterized genes, most of which have TF binding domains, such as Zn fingers
## Gene prioritization via genomic data fusion

### Gene prioritization based on known or putative involvement in cardiac arrhythmia

<table>
<thead>
<tr>
<th>Gene</th>
<th>Known or putative involvement in cardiac arrhythmia</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCN5A</td>
<td>Sodium channel, voltage-gated, type V, included in the training set</td>
</tr>
<tr>
<td>CACNA1D</td>
<td>Voltage-dependent L-type calcium channel alpha-1D subunit, belongs to the same family of channels as CACNA1C, which is included in the training set</td>
</tr>
<tr>
<td>CAV1.3</td>
<td>Cav1.3-knockout mice suffer from disturbed atrio-ventricular conduction and abnormal contractile function</td>
</tr>
<tr>
<td>ITPR1 (IP3R)</td>
<td>Inositol 1,4,5-trisphosphate receptor which regulates Ca2+ homeostasis in the heart. Defective cellular calcium is widely recognized as a significant pathophysiological event in the contractile dysfunction of the failing heart</td>
</tr>
<tr>
<td>TRPC1</td>
<td>An ion channel involved in regulating the flux of small cations (including sodium and calcium), is also strongly expressed in cardiac myocytes</td>
</tr>
<tr>
<td>SCN10A</td>
<td>Sodium channel, voltage-gated, type X</td>
</tr>
<tr>
<td>SCN11A</td>
<td>Sodium channel, voltage-gated, type XI</td>
</tr>
<tr>
<td>CICN2</td>
<td>Chloride (Cl) ion channel mutated in epilepsy, many channels in the central nervous system are also contributing to cardiac electrophysiology</td>
</tr>
<tr>
<td>ENSG00000172139</td>
<td>Uncharacterized gene with cation channel activity</td>
</tr>
<tr>
<td>ENSG00000131388, ENSG00000144712 and ENSG00000168016</td>
<td>Uncharacterized genes with an ankyrin domain, which is over-represented in the training set</td>
</tr>
</tbody>
</table>

### Gene prioritization based on known or putative involvement in cardiomyopathy

<table>
<thead>
<tr>
<th>Gene</th>
<th>Known or putative involvement in cardiomyopathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAV3</td>
<td>Caveolin-3 Knock-out Mice Develop a Progressive Cardiomyopathy, included in the training set</td>
</tr>
<tr>
<td>MYBPC2</td>
<td>Causes hypertrophic cardiomyopathy, included in the training set</td>
</tr>
<tr>
<td>TNNC1</td>
<td>Troponin C is one of the 3 subunits that form troponin, which together with tropomyosin is responsible for the regulation of striated muscle contraction. Although not included in the training set, mutations have been found in patients with CM</td>
</tr>
<tr>
<td>CTNNB1</td>
<td>Deletion of beta-catenin in the endoderm results in multiple embryonic hearts. In cardiomyopathic hearts, CTNNB1 accumulates in intercalated disks</td>
</tr>
<tr>
<td>RAF1</td>
<td>Mechanical stretch is an initial factor for cardiac hypertrophy by activating second messengers such as Raf-1 kinase</td>
</tr>
<tr>
<td>PKCD</td>
<td>Loss of PKC-delta alters cardiac metabolism</td>
</tr>
<tr>
<td>FLNB</td>
<td>Filamin plays an important role in intracellular and intercellular linkages in cardiac muscle</td>
</tr>
<tr>
<td>DAG1</td>
<td>Through dystrophin and actin interactions, the dystrophin-associated complex (α- and β-dystroglycans, α-, β-, γ- and δ-sarcoglycans, CAV3, syntrophin, and dystrobrevin) provide stability to the sarcomere and transmit force to the extracellular matrix</td>
</tr>
<tr>
<td>LAMR1</td>
<td>LAMR1 functional retroposon in mice causes right ventricular dysplasia, which is a hereditary cardiomyopathy that causes sudden death in the young</td>
</tr>
<tr>
<td>MYLK</td>
<td>Myosin light chain kinase</td>
</tr>
</tbody>
</table>

---

1. Aerts et al. 2021
2. Aerts et al. 2021
3. Aerts et al. 2021
4. Aerts et al. 2021
5. Aerts et al. 2021
6. Aerts et al. 2021
7. Aerts et al. 2021
8. Aerts et al. 2021
9. Aerts et al. 2021
10. Aerts et al. 2021
11. Aerts et al. 2021
12. Aerts et al. 2021
13. Aerts et al. 2021
14. Aerts et al. 2021
15. Aerts et al. 2021
16. Aerts et al. 2021
17. Aerts et al. 2021
18. Aerts et al. 2021
19. Aerts et al. 2021
20. Aerts et al. 2021
Table, listing a number of putative disease candidates, involved in congenital heart disease (CHD), obtained after prioritization by ENDEAVOUR of all 1,048 test genes on chromosome 3 and a selected list of the top 5% best ranked genes. For each of these candidates, a short description of their possible function is enclosed, revealing that each of them is indeed somehow involved in cardiovascular development, arrhythmic dysfunction, or cardiomyocyte function respectively.

METHODS FOR CHD: The following CHD training genes were selected: NKX2.5, GATA4, TBX5, TBX1, JAGGED1, PROSIT240, CFC1, FOG2, CRELD1, PTPN11, SEMA3E,21,22. Ensembl IDs from these genes are available upon request. The data sources, which were used to extract the training information for the prioritizations are: EnsemblEst expression, Microarray Expression, Gene Ontology, InterPro, cis-regulatory modules, transcription motifs, Blast, Literature, KEGG and BIND. The following over-represented ("specific") attributes for the CHD training set were used: heart and cardiovascular system for the EnsemblEst expression data source; heart, development, regulation of transcription factor activity and DNA binding for Gene Ontology; heart, transcription, development, DNA binding for the Literature data source; T-box and Brachury transcription factor, GATA-type transcription activator, Zn-finger and EGF-like domains for the InterPro source; ARNT/SP1/CAC_BP/LYF1/AP-2 cis-regulatory module for the cis-regulatory elements’ data source; CREB, AP1, AP2, ATF3/4, GC sequence motifs for the Transcriptional motifs’ source; and high average expression in heart, bronchial epithelial cells, and cardiac myocytes for the microarray expression data source.

METHODS FOR AR: The following AR training genes were selected: RYR-2, KCNQ1, ANKYRI, KCNH2, KCNE2, KCNE1, KCNJ2, PRKAG2, SCN5A, CASQ2. Ensembl IDs from these genes are available upon request. The data sources, which were used to extract the training information for the prioritizations are: EnsemblEst expression, Microarray Expression, Gene Ontology, InterPro, cis-regulatory modules, transcription motifs, Blast, Literature, KEGG and BIND. The following over-represented ("specific") attributes for the AR training set were used: epidermis, skin, myocardium, trabecular meshwork, heart for the EnsemblEst expression data source; regulation of heart rate, muscle contraction, cation channel activity, circulation, cell motility, metal ion transport for Gene Ontology; heart, transcription, development, potassium channel, death, heart, conduction, voltage for the Literature data source; voltage-gated potassium and different types of cation channels for the InterPro source; MEF2/OCT1/CAC_BP/CEBP/E2A (e.g. MEF2 is well known to be involved in cardiac muscle development28) for the cis-regulatory modules; high average expression in heart, endothelial, whole-blood, thyroid, and skeletal muscle for the microarray expression data source.

METHODS FOR CM: The following CM training genes were selected: dystrophin, tafazzin, actin, desmin, d-sarcoglycan, troponin T, β-myosin heavy chain, α-tropomyosin, laminin A/C, troponin I, myosin-binding protein C, myosin essential light chain, myosin regulatory light chain, titin, dystrobrevin14,26. Ensembl IDs from these genes are available upon request. The data sources, which were used to extract the training information for the prioritizations are: EnsemblEst expression, Microarray Expression, Gene Ontology, InterPro, cis-regulatory modules, transcription motifs, Blast, Literature, KEGG and BIND. The following over-represented ("specific") attributes for the CM training set were used: omentum, heart, muscle, myocardium, musculoskeletal for the EnsemblEst expression data source; cell motility, (striated) muscle development and contraction, kinesin complex, cytoskeleton, myofibril, sarcomere, actin for Gene Ontology; muscle, skeletal, myosin, troponin, actin, tropomyosin, filament for the Literature data source; protein domains such as troponin, tafazzin, tropomyosin, KCNQ1 voltage-gated potassium channel, synuclein, sarcoglycan complex, PPARK motif, intermediate filament, sulphonylurea receptor for the InterPro source; the T3R/Myogenin/CAC_BP/SP1/MEF2 (Myogenin and MEF2 are well known to be involved...
in cardiac muscle development (25) cis-regulatory module for the cis-regulatory elements’ data source; NFAT, MEF2, myogenin, STAT3 binding sites (NFAT is a signaling molecule known to be involved in cardiac hypertrophy (27)) for the transcriptional motifs’ source.

REFERENCES