

# Identification of a peripheral blood gene signature predicting aortic valve calcification

Donal Macgrogan, Beatriz Martínez-Poveda, Jean-Pierre Desvignes, Leticia Fernandez-Friera, Manuel José Gomez, Eduardo Gil Vilariño, Sergio Callejas Alejano, Pablo García-Pavía, Jorge Solis, Joaquín Lucena, et al.

### ▶ To cite this version:

Donal Macgrogan, Beatriz Martínez-Poveda, Jean-Pierre Desvignes, Leticia Fernandez-Friera, Manuel José Gomez, et al.. Identification of a peripheral blood gene signature predicting aortic valve calcification. Physiological Genomics, 2020, 52 (12), pp.563-574. 10.1152/physiolgenomics.00034.2020. hal-02965646

# HAL Id: hal-02965646 https://amu.hal.science/hal-02965646

Submitted on 28 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 2	<u>MS PG-00034-2020 revised</u> : Identification of a peripheral blood gene signature predicting
2	aortic valve calculation
) 1	Running title: A gene signature predicting aortic valve calcification
+ 5	Supplemental Material and Suppl. Tables available at:
<b>)</b> 7	URL: https://figshare.com/s/d9a68e0d6faa67e4c6c3 and https://figshare.com/s/4357d56025cb3380a0df
5 )	DOI: 10.6084/m9.figshare.12086712 and 10.6084/m9.figshare.12086724
)	
	Donal MacGrogan <sup>1,2,*</sup> , Beatriz Martínez Poveda <sup>1,2</sup> , Jean-Pierre Desvignes <sup>3</sup> , Leticia Fernandez-Friera <sup>2,4,5</sup> ,
	Manuel José Gomez <sup>6</sup> , Eduardo Gil Vilariño <sup>7</sup> , Sergio Callejas Alejano <sup>7</sup> , Pablo Garcia-Pavia <sup>2,8,9</sup> , Jorge
	Solis <sup>10</sup> , Joaquín Lucena <sup>11</sup> , David Salgado <sup>3</sup> , Gwenaelle Collod-Béroud <sup>3</sup> , Emilie Faure <sup>3</sup> , Alexis Théron <sup>12</sup> ,
	Julia Torrents <sup>13</sup> , Jean-François Avierinos <sup>3,12</sup> , Lorena Montes <sup>14</sup> , Ana Dopazo <sup>7</sup> , Valentín Fuster <sup>15,16</sup> , Borja
	Ibañez <sup>2,17</sup> , Fátima Sánchez-Cabo <sup>6</sup> , Stephane Zaffran <sup>3</sup> & José Luis de la Pompa <sup>1,2,*</sup>
	<sup>1</sup> Intercellular Signaling in Cardiovascular Development & Disease laboratory, Centro Nacional de
	Investigaciones Cardiovasculares Carlos III (CNIC), Melchor Fernández Almagro 3, 28029 Madrid,
	SPAIN
	<sup>2</sup> Ciber de Enfermedades Cardiovasculares, Instituto de Salud Carlos III, Melchor Fernández Almagro 3,
	28029 Madrid, SPAIN <sup>3</sup> Aix Marsaille University MMC, INSERM U1251, Marsaille, ERANCE
	<sup>4</sup> Translational Laboratory for Cardiovascular Imaging and Therapy CNIC SPAIN
	<sup>5</sup> HM Hospitales-Centro Integral de Enfermedades Cardiovasculares, Madrid, SPAIN
	<sup>6</sup> Bioinformatics Unit, CNIC, SPAIN
	<sup>7</sup> Genomics Unit, CNIC, SPAIN
	<sup>8</sup> Departmento de Cardiología, Hospital Universitario Puerta de Hierro, Madrid, SPAIN
	<sup>9</sup> Universidad Francisco de Vitoria, Pozuelo de Alarcón, Madrid, SPAIN
	<sup>10</sup> Departmento of Cardiología, Hospital Universitario Doce de Octubre, Madrid, SPAIN
	Servilla SPAIN
	<sup>12</sup> Service de Cardiologie. Hôpital de la Timone. 13005. Marseille. FRANCE
	<sup>13</sup> Service d'anatomie et Cytologie Pathologiques, Hôpital de la Timone, 13005, Marseille, FRANCE
	<sup>14</sup> Hospital Clínico San Carlos, Madrid, SPAIN
	<sup>15</sup> Cardiovascular Imaging and Population Studies Laboratory, CNIC, SPAIN
	<sup>16</sup> Cardiology Department, Icahn School of Medicine at Mount Sinai, New York, New York, USA
	<sup>17</sup> IIS-Fundación Jiménez Díaz Hospital Universitario, Madrid, SPAIN
	* Corresponding authors
	Correspondence:
	Intercellular Signaling in Cardiovascular Development & Disease laboratory
	Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC)
	Melchor Fernández Almagro 3, 28029 Madrid, SPAIN.
	Tel: +34 91 4531200

#### 47 ABSTRACT

48

49 Calcific aortic valve disease (CAVD) is a significant cause of illness and death worldwide. 50 Identification of early predictive markers could help optimize patient management. RNA-51 sequencing was carried out on human fetal aortic valves at gestational weeks 9, 13, and 22, and 52 on a case-control study with adult non-calcified and calcified bicuspid and tricuspid aortic 53 valves. In dimension reduction and clustering analyses, diseased valves tended to cluster with 54 fetal valves at week 9 rather than normal adult valves, suggesting that part of the disease 55 program might be due to re-iterated developmental processes. The analysis of groups of co-56 regulated genes revealed predominant immune-metabolic signatures, including innate and 57 adaptive immune responses involving lymphocyte T cell metabolic adaptation. Cytokine and 58 chemokine signaling, cell migration, and proliferation were all increased in CAVD, whereas 59 oxidative phosphorylation and protein translation were decreased. Discrete immune-metabolic 60 gene signatures were present at fetal stages and increased in adult controls, suggesting that these 61 processes intensify throughout life and heighten in disease. Cellular stress-response and 62 neurodegeneration gene signatures were aberrantly expressed in CAVD, pointing to a 63 mechanistic link between chronic inflammation and biological ageing. Comparison of the valve 64 RNA-sequencing dataset with a case-control study of whole blood transcriptomes from 65 asymptomatic individuals with early aortic valve calcification identified a highly predictive gene 66 signature of CAVD and of moderate aortic valve calcification in overtly healthy individuals. 67 These data deepen and broaden our understanding of the molecular basis of CAVD and identify 68 a peripheral blood gene signature for the early detection of aortic valve calcification.

69

#### 70 Keywords

71 Human fetal valve; CAVD; inflammation; peripheral blood biomarker; gene signature

#### 72 INTRODUCTION

73 Calcific aortic valve disease (CAVD) is caused by calcium build-up of the aortic valve leaflet 74 and may be associated with hemodynamic obstruction, leading to different degrees of aortic 75 stenosis (AS). Untreated severe AS leads to heart failure, with valve replacement remaining the 76 only available treatment. Severe AS is preceded by aortic sclerosis, a subclinical form of CAVD 77 that develops over many years without hemodynamic consequence (6). The prevalence of aortic 78 valve disorders increases with age, with 25% of individuals >65 years with aortic sclerosis 79 progressing to AS at a rate of 2% per year (4). Other than age, the principal CAVD risk factor for 80 AS is a congenitally malformed bicuspid aortic valve (BAV), which is present in 0.5-2% of the 81 general population (31). Patients with AS and a BAV are a decade younger at diagnosis than 82 those with normal tricuspid aortic valve (TAV) morphology (2). Premature leaflet thickening and 83 calcification in BAV patients may be caused by abnormal shear stress hemodynamics.

84 CAVD shares mechanistic features with atherosclerosis (6). The current accepted model 85 implicates initial lipid oxidation, inflammatory infiltration by T-lymphocytes and macrophages, 86 cytokine and pro-fibrotic signaling, and altered extracellular matrix (ECM) secretion, leading to 87 ectopic bone formation and tissue mineralization (11, 22). The factors predicting transition from 88 the "at risk" status to aortic sclerosis and subsequently to severe AS, are not known (25). 89 Transcriptional profiling of CAVD has helped to identify hundreds of genes differentially 90 expressed between calcified and normal valves (8). Multi-omics approaches combining 91 transcriptomics and proteomics have provided a high-resolution expression atlas and established 92 potential drivers of disease progression (30).

Despite this progress, thus far no medical therapy has been proven to be effective in preventing progression to severe disease, with randomized clinical trials demonstrating the lack of effect of statins on disease progression (36). It has been suggested that statin treatment would have to begin at an earlier disease stage to have a beneficial effect. Current imaging modalities, especially echocardiography, remain the gold standard for diagnosis of valve disease but are

98 mainly performed when symptoms appear. Non-contrast computed tomography may have 99 significant predictive value in both asymptomatic and symptomatic patients, but remains a 100 source of radiation exposure and may not be available in centers with limited resources. 101 Biomarkers may complement imaging modalities to detect early changes, especially early 102 calcification. The ideal biomarker would predict disease as well as improve prognosis, 103 subclinical disease management, and risk stratification and reduce disease-associated morbidity 104 and mortality. Already identified potential biomarkers include markers of myocardial injury, 105 cardiac mechanical stretch, inflammation, and hemostasis imbalance; however, with the 106 exception of lipoprotein(a), these are not specific to CAVD or have yet to be validated in large 107 randomized studies (32).

108 The aortic valve develops from rudimentary endocardial cushions that are subsequently 109 refined into thin leaflets well into postnatal life. This process is dependent on complex cell-ECM 110 interactions and is modulated by hemodynamics, and its spatio-temporal dysregulation results in 111 malformations and maladaptive remodeling (10). Several lines of evidence indicate that valve 112 disease at any age is accompanied by some underlying structural defect (29), suggesting that 113 latent CAVD has a developmental origin. Moreover, developmental pathways may also be re-114 activated during tissue repair processes in the disease setting. For example, molecular 115 mechanisms controlling heart valve cell differentiation are shared with cartilage, tendon, and 116 bone development (reviewed in (15), and are reactivated during endochondral ossification and 117 mature lamellar bone formation taking place in end-stage aortic valve disease (18). Remarkably, 118 only end-stage CAVD has been described at the molecular level, whereas the molecular 119 underpinnings of subclinical disease are poorly understood, due in part to difficulties in 120 accessing biological material from early lesions.

121 To gain mechanistic insight, we performed an RNA-seq analysis of human fetal and adult 122 control and calcified bicuspid (cBAV) and tricuspid valves (cTAV). K-means clustering was 123 used to identify dynamically regulated processes taking place simultaneously at crucial stages of 124 human valve development and in diseased valves. The prevailing gene signatures were 125 associated with immune-metabolic processes and heretofore uncharacterized pathways linked to 126 cellular stress and ageing. A search for genes shared between our gene expression datasets and 127 datasets generated from asymptomatic subjects from the Progression of Early Subclinical 128 Atherosclerosis (PESA) cohort (5) with aortic valve calcification and without AS identified 129 marker genes in peripheral blood of asymptomatic individuals. When combined in quantitative 130 RT-PCR assays, these marker genes showed good sensitivity and specificity for predicting 131 CAVD at a subclinical stage.

#### 132 **RESULTS**

#### 133 Diseased valves are transcriptionally more similar to developing valves than to controls 134 Genome-wide transcriptome profiling was carried out on normal adult control valves 135 (n=8) and severely stenotic calcified BAV (cBAV; n=5) and TAV (cTAV; n=7) valves, as well 136 as on fetal aortic values at gestational weeks 9 (n=4), 12-13, (n=3) and 22 (n=2) (Supplementary 137 Table S1, see https://figshare.com/s/4357d56025cb3380a0df). These developmental timepoints correspond to mouse embryonic days (E) 13.5-17.5 (13), during which valves are remodeled by 138 139 morphogenic processes leading to thinning, elongation, and formation of multi-laminar valve 140 structure. To integrate data across studies and generalize our findings, we conducted a meta-141 analysis with publicly available expression data from normal adult control (n=8), cBAVs (n=10), 142 and cTAVs (n=9) valves (8) (Figure 1A). After correcting for batch effects, Principal Component 143 Analysis (PCA) indicated that sample differences did not correlate with batch origin 144 (Supplementary Fig. S1A), but rather depended on disease or developmental stages (Figure 1B). 145 The first of the PCA components (PC1) separated developing and adult valve samples according 146 to disease severity, ranging from cTAV (right hand side) to controls (left hand side). 147 Interestingly, fetal valves tended to cluster with diseased adult valves. There were two control 148 samples (both from heart transplants) and one sample of embryonic samples at 9 weeks showing 149 more heterogeneity than the rest according to the second component (y-axis) but still the range of 150 the axis (-0.6 to 0.4) is very small, confirming the reproducibility of the samples in terms of 151 global transcriptome. In the second PCA component (PC2), the adult control and disease samples 152 were separated from fetal weeks 9 and 13 samples (Figure 1B). Therefore, and considering that the first PCA dimension explains 15.24% of the variance, while the second dimension explains 153 154 only 7.73%, adult disease samples appeared to be more closely related to fetal than to control 155 samples, suggesting that some of the disease pathogenesis can be ascribed to the recapitulation of 156 developmental processes.

- 157 <u>To further support this finding, we used Euclidean distances between overall gene</u>
- 158 expression profiles to perform hierarchical clustering (HC) of the samples (Figure 1C). Similar

159 to what was observed in the PCA, disease samples cluster further apart from controls and appear

- 160 <u>under the same branch with embryonic samples at 9 and 13 weeks.</u>
- 161

#### 162 Clustering identifies dynamic gene expression across disease and developmental states

163 Differential expression analysis was initially carried out between each pair of conditions of 164 interest (Supplementary Table S2, see https://figshare.com/s/4357d56025cb3380a0df

165 ; 9 pairwise comparisons). However, pair-wise comparisons were not adequate to explore 166 dynamic and coordinated gene expression profiles across different disease and developmental 167 states. K-means analysis computing n=50 yielded dynamic expression profiles across the 168 different conditions (Supplementary Fig. S1B, see 169 https://figshare.com/s/d9a68e0d6faa67e4c6c3). Co-expressed genes (ie. genes undergoing 170 similar disease/stage status expression changes) likely participate in the same process, and genes 171 that function in the same process are regulated coordinately (20). There were 6 broad combinations of expression patterns: 1) genes down-regulated in developing valves and up-172 173 regulated in control adult valves (e.g. C34); 2) genes down-regulated in control and up-regulated 174 in diseased valves (e.g. C7); 3) genes up-regulated in developing valves, down-regulated in adult 175 control valves, and up-regulated in diseased valves (e.g. C45); 4) genes up-regulated in 176 developing valves and down-regulated in adult control and disease valves (e.g. C37); 5) genes 177 up-regulated in adult control valves and down-regulated in diseased valves (e.g. C16); and 6) 178 genes down-regulated in developing valves, up-regulated in adult control valves, and down-179 diseased C40) regulated in valves (e.g. (Supplementary Fig. S1B. see 180 https://figshare.com/s/d9a68e0d6faa67e4c6c3). Thus, part of the disease process involves a gene 181 signature present during fetal valve development.

#### 183 Comparative functional analysis identifies multiple interconnected pathways

184 Functional enrichment analysis of clusters with IPA identified significant associations to 185 canonical pathways in 28 clusters (Figure 2). Subsequent unsupervised hierarchical clustering of 186 clusters, based on the fraction of shared canonical pathways, identified 6 metaclusters (MC), 187 termed MC1-MC6 (Figure 2). Cluster 10 did not overlap with the others in terms of functional 188 enrichment, and was not further considered. Generally, genes highly expressed in adult disease 189 showed low expression at gestational week 13 and vice-versa; however, there were exceptions 190 (clusters 9, 12, 23, and 45; Figure 2). Although metaclusters were defined as non-overlapping 191 collections of gene clusters, enriched functions could be shared between several meta-clusters. 192 For example, canonical pathways associated to clusters 46 and 32 were shared by MC1 and 193 MC2; those associated to cluster 12 were shared by MC2, MC3, MC4, and MC6; those to 194 clusters 12 to 45 were shared by MC2, MC3, and MC4; and those canonical pathways enriched 195 in clusters 5 and 45 were shared by MC4 and MC6 (Figure 2). Therefore, the biological 196 functions represented in the different MCs are interconnected.

197

#### 198 Innate and adaptive immune processes are present in valve disease and development

199 The collections of genes associated to metaclusters MC1-MC6 were reanalysed with IPA. 200 Enriched canonical pathways associated to each MC are summarized in Figure 3. MC3 was 201 associated to 168 pathways related to innate-adaptive immune processes (Supplementary Fig. S2, 202 see https://figshare.com/s/d9a68e0d6faa67e4c6c3). The top 5 pathways were Th1 and Th2 203 Activation Pathway, TREM1 Signaling, Communication between Innate and Adaptive Immune 204 Cells, Phagosome Formation and Dendritic Cell Maturation (Figure 3; Supplementary Fig. S3, 205 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3. see see 206 https://figshare.com/s/4357d56025cb3380a0df). The remaining 163 enriched biological functions were also involved in critical immune-inflammatory cellular processes (Supplementary 207 208 Fig. S2, see https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see

https://figshare.com/s/4357d56025cb3380a0df). Among the nine clusters combined into MC3, 209 210 five of them had expression profiles characterised by increased gene expression in all adult 211 samples, both control and CAVD (clusters 22, 9, 34, 36 and 2), whereas four clusters had 212 increased expression in CAVD only (clusters 49, 44, 7, and 35; Figure 3). Discrete immune cell 213 gene signatures were present at week 9 (clusters 44 and 35), week 13 (cluster 9), and week 22 214 (clusters 49 and 22) of development (Figure 3), but involved relatively fewer pathways than 215 those found in control or calcified adult valves. The presence of immune and pro-inflammatory 216 gene signatures during fetal development and in apparently healthy valves is consistent with the 217 progressive nature of the calcified valve disease pathology.

218

#### 219 Cell migration and DNA replicative pathways are re-activated in CAVD

220 Metaclusters MC4 and MC6 were enriched in genes belonging to 202 and 23 IPA pathways, 221 S4 respectively (Supplementary Fig. and S6A, see 222 https://figshare.com/s/d9a68e0d6faa67e4c6c3; S3, Supplementary Table see 223 https://figshare.com/s/4357d56025cb3380a0df). In general, genes associated to these clusters 224 have increased expression during development (notably at week 13), decreased in adulthood and 225 re-activated in CAVD. Therefore, of the 7 clusters associated with these two MCs, six were 226 associated with more elevated gene expression in the diseased state (clusters 25, 23, 5, 45, 17, 227 and 21), whereas only one was associated with lower expression (cluster 40; Figure 3). The top 228 five pathways in MC4 were Actin Cytoskeleton Signaling, Molecular Mechanisms of Cancer, 229 Axonal Guidance Signaling, Epithelial Adherens Junction Signaling and Integrin Signaling 230 3. S5, S3, (Figure Supplementary Fig. Supplementary Table see 231 https://figshare.com/s/4357d56025cb3380a0df). The remaining 197 IPA pathways, represented 232 components of fundamental cell and developmental signaling pathways (Supplementary Fig. S4, 233 see https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 234 https://figshare.com/s/4357d56025cb3380a0df). The top five pathways in MC6 were Role of 235 CHK Proteins in Cell Cycle Checkpoint Control, Sumoylation Pathway, Cell Cycle Control of 236 Chromosomal Replication, Cell Cycle: G2/M DNA Damage Checkpoint Regulation, Role of 237 DNA Damage Response (Figure 3, Supplementary Fig. S6B, BRCA1 in see 238 https://figshare.com/s/d9a68e0d6faa67e4c6c3, Table S3, Supplementary see 239 https://figshare.com/s/4357d56025cb3380a0df). The remaining 18 IPA pathways were involved 240 similar in (Supplementary Fig. S6A, processes see 241 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 242 https://figshare.com/s/4357d56025cb3380a0df). Cellular proliferation and migration may be 243 part of the ongoing immune-inflammatory response associated with processes represented in 244 MC3.

245

#### 246 Decreased OXPHOS and ribosomal biogenesis in CAVD

247 Genes in MC2 were significantly associated to 83 IPA pathways (Supplementary Fig. S7, see 248 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see https://figshare.com/s/4357d56025cb3380a0df), the top 5 of which were Mitochondrial 249 250 Dysfunction, EIF2 Signaling, Oxidative Phosphorylation, Sirtuin Signaling Pathway 251 (Supplementary Fig. S8A-C), Regulation of eIF4 and p70S6K Signaling (Supplementary Fig. 252 S8D,E see https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 253 https://figshare.com/s/4357d56025cb3380a0df). The remaining 78 pathways in MC2 were also 254 implicated in energy production and protein metabolism (Supplementary Fig. S7, see 255 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 256 https://figshare.com/s/4357d56025cb3380a0df), showing strong enrichment in genes for nuclear 257 encoded subunits of the electron transport chain complexes and ribosomal biogenesis 258 (Supplementary Fig. S8A-C, see https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary 259 Table S3, see https://figshare.com/s/4357d56025cb3380a0df). Other gene signatures were AMP-260 activated protein kinase (AMPK), ER stress (Unfolded protein response), oxidative stress

261 response (NRF2-mediated Oxidative Stress Response), cellular longevity (Telomerase 262 Signaling), and neurodegenerative disease (Huntington's Disease Signaling; Supplementary Fig. 263 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3. S7, see see 264 https://figshare.com/s/4357d56025cb3380a0df), suggesting the modulation of cellular stress 265 responses, and age-related disease gene signatures. Of the nine clusters in MC2, six had, on 266 average, decreased expression in disease conditions (clusters 24, 31, 32, 16, 41, and 1), while 267 one had increased expression in the diseased state (cluster 12) and three were unchanged 268 (clusters 31, 26, and 37; Figure 3). In general, MC2 transcripts had peak expression at week 13, 269 with lower expression at the other developmental stages (Figure 3), suggesting that metabolic 270 processes are dynamically regulated during aortic valve development. Overall, this analysis 271 indicates that CAVD is associated with decreased OXPHOS metabolism and altered proteostasis.

272

#### 273 Stress response gene signatures are altered in CAVD

274 MC1 and MC5 did not overlap significantly with any of the other MCs but included similar 275 biological processes. Genes in MC1 were associated to 7 pathways (Figure 3; Supplementary 276 Fig. S9A, see https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 277 https://figshare.com/s/4357d56025cb3380a0df), the top 5 being Sphingosine-1-phosphate 278 Signaling, Endocannabinoid Developing Neuron Pathway, Amyloid Processing, Huntington's 279 Disease Signaling, Dermatan Sulfate Degradation (Figure 3; Supplementary Fig. S9B, see 280 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3. see 281 https://figshare.com/s/4357d56025cb3380a0df). The expression of MC1 genes was relatively 282 increased in adult controls and decreased in CAVD (Figure 3). MC5 was enriched in a single 283 pathway, Unfolded protein response (Figure 3; Supplementary Fig. S9C, see 284 https://figshare.com/s/d9a68e0d6faa67e4c6c3; Supplementary Table S3, see 285 https://figshare.com/s/4357d56025cb3380a0df). MC5 genes were expressed at decreased levels 286 in controls and increased in CAVD (Figure 3). These data suggest that chronic immuneinflammatory processes are linked to cellular stress and age-related disease in calcific aorticvalve pathology.

289

290

#### 291 Identification of circulating biomarkers predicting CAVD

292 To identify potential biomarkers useful for predicting CAVD, a sub-set of genes from the aortic 293 valve study was compared with whole blood RNA-Seq data from asymptomatic individuals with 294 aortic valve calcification, detected by non-contrast computed tomography in the *Progression of* 295 Early Subclinical Atherosclerosis (PESA) study (5). PESA participants undergo simultaneous 296 cardiac calcium scoring and blood sampling. Out of the 455 PESA participants for which RNA-297 seq data was available in visit 2 (ie. three years after recruitment), 59 had evidence of valve 298 calcification and 52 of theses had sufficient integrity for qPCR analysis. None of them had 299 valvular stenosis/regurgitation, evidencing an early stage of the disease. Their transcriptome was 300 compared to the transcriptome of PESA individuals without valvular calcification (n=395). From 301 the whole set of genes identified as differentially expressed (p < 0.05, Supplementary Table S5, 302 see https://figshare.com/s/4357d56025cb3380a0df), six were selected that showed a disease-303 specific pattern in the RNA-seq data of human aortic valves (see Methods for selection criteria; 304 Supplementary Fig. S1, see https://figshare.com/s/d9a68e0d6faa67e4c6c3). Five of these genes 305 were related to adaptive innate and adaptive responses (CD28, ITK, PAG1, NLRC5, and ASCC3) 306 and a sixth was related to the WNT pathway (LINS1) (Supplementary Table S6, see 307 https://figshare.com/s/4357d56025cb3380a0df). The expression of the six genes was measured in 308 whole blood (WB) on 104 PESA participants [52 cases: asymptomatic individuals with 309 calcification of the aortic valve (PESA-SD); 52 controls: individuals without calcification 310 (PESA-CTL)] and 12 patients with CAVD recruited at Monteprincipe Hospital (WB-CAVD) by 311 <u>qRT-PCR</u> and correlated with the degree of aortic valve calcification (Figure 4A, Supplementary 312 Table S4, see https://figshare.com/s/4357d56025cb3380a0df). Considering all conditions, the 6

genes showed some degree of correlation when considered individually ( $P \le 0.27$ ; Figure 4A, Supplementary Table S7, see <u>https://figshare.com/s/4357d56025cb3380a0df</u>). Given that, *NLRC5* and *PAG1* were also associated with total calcification of the coronaries (Supplementary Fig. S10A, see <u>https://figshare.com/s/d9a68e0d6faa67e4c6c3</u>, Supplementary Table S7, see <u>https://figshare.com/s/4357d56025cb3380a0df</u>) and were therefore excluded as potential specific markers of aortic valve calcification.

To determine the utility of *ASCC3*, *ITK*, *CD28* and *LINS1* as potential markers of valve calcification, a ROC curve was built to represent the predictive value of the model to classify calcific disease patients (WB-CAVD) vs PESA-SD individuals and vs PESA-CTL, yielding areas under the curve (AUCs) of 0.97 and 0.91 respectively (Figure 4B). An additional model was built to assess the power of the gene signature to distinguish between individuals without calcification (PESA-CTL) from those with subclinical disease (PESA-SD; AUC=0.71).

325 <u>This model was then validated in whole blood samples of an independent cohort of</u> 326 <u>CAVD patients (n=16), together with samples from 25 of the 52 PESA-SD individuals and 35 of</u>

327 the 52 PESA-CTL individuals taken three years after visit 2 (Figure 4C). The model achieved

328 AUCs of 0.90 and 0.88 to discriminate CAVD patients from PESA-SD and PESA-CTL

329 individuals, respectively. The AUC of the model to distinguish PESA-SD from PESA-CTL was

330 <u>again 0.71 (Figure 4C)</u>. Thus, our model identifies a gene expression signature in blood that is

highly predictive of valve disease and of aortic valve calcification levels in otherwise healthy

individuals.

#### 334 **DISCUSSION**

335 This study provides, for the first time, a comprehensive molecular profile of shared and unique 336 cellular activities taking place in aortic valve development and CAVD. The transcriptome of 337 adult disease samples appeared to be more closely related to fetal week 9 than to control samples 338 or weeks 13 and 22, suggesting that some of the disease pathogenesis might be ascribed to the 339 recapitulation of processes taking place in early development. The overall cBAV and cTAV gene 340 expression profiles were essentially identical, confirming previous findings suggesting that valve 341 calcification depends on the same biological pathways, regardless of valve morphology (8, 26). 342 Therefore, the mechanisms leading to BAV formation are independent of those leading to BAV 343 calcification; in other words, the cause of the embryonic defect is distinct from the resulting 344 calcification process.

345 Functional analysis provided compelling evidence that innate and adaptive immune 346 interactions take place in the aortic valve throughout life (Supplementary Fig. S10B 13, see 347 https://figshare.com/s/d9a68e0d6faa67e4c6c3). The fetal immune cell gene signature in MC3 348 was notably simpler than in adult control or CAVD samples, with fewer immune cell gene 349 expression signatures present at gestational weeks 9 (3/9 clusters), 13 (1/9 clusters), or 22 (3/9 350 clusters) than in adult controls (5/9 clusters) or in CAVD (9/9 clusters), suggesting that aortic 351 valve immune cell enrichment in the aortic valve increases over time. Moreover, the presence of 352 only one immune cell gene signature at 13 weeks (cluster 9), compared with three at 9 and 22 353 weeks of gestation, implies that the immune landscape is dynamically regulated in developing 354 valves. These observations are consistent with previous findings of multiphasic changes in shear 355 stress-response, cell-ECM interaction and adaptive immunity pathways from week 14 to week 22 356 of human aortic valve development (7).

The predominant T cell gene signature in the RNA-seq dataset strengthens the notion that unresolved immune-inflammatory response is a major CAVD component (17). This association is well established, with early immunohistological findings indicating the presence of T

360 lymphocytes in immune infiltrates (23, 24). More recent studies have revealed the existence of T 361 cell infiltration in BAV and TAV calcific disease, consisting of non-specific chemokine driven 362 polyclonal T cell recruitment, and memory-effector CD8+ T cell clonal expansions (34, 35), 363 supporting the interpretation of an antigen-induced immune response. Moreover, this antigen-364 driven response involved circulating expanded CD8+ and, to a lesser extent, CD4+ T cell clones 365 between circulating blood and valve tissues suggesting a systemic disease (34). The gene 366 expression signatures of several other immune-competent cells, including B-lymphocytes, 367 macrophages, and dendritic cells were also present, confirming previous characterization of 368 cellular infiltrates in CAVD (33). Together with a body of literature suggesting a tight 369 inflammatory cell-valve interstitial cell interaction in CAVD (11, 17), our data are consistent 370 with the notion that unresolved inflammation ultimately drives myo-fibroblast and osteoblast-371 like differentiation and leads to heterotopic bone formation. Therefore, targeted therapeutic 372 immunosuppression of calcification might be approachable by better characterizing the various 373 infiltrating immune-cell subtypes.

374 Functional analysis revealed that OXPHOS metabolism was decreased in CAVD, with 375 coordinated expression reduction of multiple nuclear genes encoding subunits of mitochondrial 376 electron transport chain complexes I, III, IV and V. These changes might be part of ongoing 377 adaptive metabolic reprogramming during T lymphocyte activation (27), when upon 378 encountering an antigen, T cells proliferate and acquire new functions, switching energy 379 metabolism from OXPHOS to glycolysis. However, both increased and decreased expression of 380 glycolytic pathway components were present, suggesting that decreased OXPHOS metabolism 381 cannot be ascribed solely to T cell activation. Moreover, adult aortic valves showed markedly 382 reduced transcript levels of multiple subunits of the small and large protein ribosomal complexes 383 (RPS and RPL respectively) and eukaryotic translation initiation factors (EIF1-4), possibly 384 reflecting co-ordinate dysregulation of OXPHOS activity and ribosomal biogenesis (19). This 385 occurred in uncalcified adult valves as well (clusters 31, 37, and 12), suggesting declining

function with age. In contrast signatures for cell death and apoptosis (for example in MC3) and, energy-demanding cell processes such as proliferation, DNA repair, and cell movement (in MC4 and MC6) were increased, possibly reflecting cellular heterogeneity within the calcified valve.

389 Decreased mitochondrial metabolism and protein translation may be a manifestation of 390 altered physiological stress responses that occur during ageing (30), which is also related to the 391 inability to maintain adaptability to stress and re-establish homeostasis in response to 392 environmental perturbations. Thus, multiple stress-response gene signatures were present in 393 MC2 and to a lesser extent in MC1 and MC5 (Supplementary Fig. S10B, see 394 https://figshare.com/s/d9a68e0d6faa67e4c6c3). These included SIRT, a highly conserved family 395 of NAD-dependent deacetylases, linking transcriptional regulation to mitochondrial function, 396 stress resistance, metabolism, cell survival and organismal longevity (3) and mTOR (Mammalian 397 target of rapamycin), a nutrient sensing system which regulates protein synthesis, to promote cell 398 growth and proliferation (19). Moreover, deficiency of the unfolded protein response (UPR) 399 pathway, which is activated by misfolding of newly synthesized proteins as part of the ER stress 400 response (9) contributes to inflammation, disease and ageing. Gene signatures including amyloid 401 processing and Huntington's disease signaling were also featured implying shared calcific valve 402 neurodegenerative disease mechanisms S10B, and (Supplementary Fig. see 403 https://figshare.com/s/d9a68e0d6faa67e4c6c3). Therefore, inflammation chronic drives 404 metabolic stress and biological ageing in CAVD.

Given the phenomenon of cellular trafficking of T cells between the diseased value tissue, and circulating blood (34), we reasoned that aortic value and circulating blood from CAVD patients share common genes that would be absent in subclinical individuals. In the PESA study we identified a subpopulation of asymptomatic individuals in this cohort with > 1value calcification detected by imaging. Therefore, we searched for circulating biomarkers by comparing the CAVD transcriptomes and whole blood from datasets of the subclinical PESA study population, and identified six differentially expressed genes, four of which were found to 412 predict subclinical and overt valve disease. Therefore *ASCC3*, *ITK*, *CD28*, and *LINS* were 413 specific to aortic valve calcification, whereas *PAG1* and *NLRC5* were also associated with 414 coronary artery calcification, suggesting that valve and vascular calcification proceed by shared 415 and separate mechanisms. These mechanisms have yet to be examined, but noteworthy is that 416 five of six genes belong to T cell activation and NF-kB pathways. This gene signature could be 417 used in combination with high-resolution imaging for the early diagnosis of valve calcification.

- 418 <u>Study Limitations</u>. A limitation of our study is the low sample numbers especially for
- 419 week 22 samples (n=2). Investigational studies involving human fetal issue are limited by ethical
- 420 and legal considerations on medical abortion. Nevertheless, our bioinformatics analysis showed
- 421 that the week 22 samples clustered together and separated well from the other groups by PCA
- 422 and we are confident of the robustness of the results involving this group. Moreover, differential
- 423 expression analysis was performed using limma, a statistical method developed ad-hoc for RNA-
- 424 seq data with low sample size (28).
- 425
- 426
- 427
- 428

#### 429 METHODS

#### 430 Study population and aortic valve collection.

431 Fetuses. Aortic valves from human fetuses were obtained from electively- (weeks 9; n=4 and 432 week 12/13; n=3) or medically (week 22; n=2) -terminated pregnancies after written informed 433 consent in concordance with French legislation (PFS14-011) and prior protocol approval from 434 the "Agence biomédecine" Table S1. de la (Supplementary see 435 https://figshare.com/s/4357d56025cb3380a0df). Aortic valves were recognized by their 436 anatomical landmarks under the microscope and leaflets were isolated with minimal aortic wall 437 contamination Samples were immediately flash frozen in liquid nitrogen. The investigation 438 conformed to the principles outlined in the Declaration of Helsinki.

439 Patients. The CAVD study population consisted of BAV (n=5) or TAV (n=7) patients referred 440 for aortic valve or aortic root surgery at Monteprincipe Hospital and Hospital Clínico San Carlos 441 (Madrid, Spain; Supplementary Table S1, see https://figshare.com/s/4357d56025cb3380a0df). 442 Patients were prospectively recruited between December 2009 and February 2013. Patients with 443 a history of rheumatic heart disease, infective endocarditis, or connective tissue disorders were 444 excluded. To determine the morphology and functional state of the aortic valve, all patients 445 underwent clinical evaluation and preoperative 2-dimensional echocardiography according to the 446 American Heart Association and the American College of Cardiology guidelines (21). The 447 aortic valve morphology and the presence of valvular degeneration, including the extent of 448 leaflet calcification, thickening, prolapse, and/or redundancy, were documented at surgery. 449 Calcification was graded as mild (1/3 leaflet area affected) or severe (2/3 leaflet area affected). 450 Immediately following surgical removal, a portion of the aortic valve leaflet was placed in sterile 451 saline buffer and stored at -80°C until processing. Control valve samples were obtained at 452 autopsy of individuals without cardiac problems who suffered a traumatic death (n=6) or at time 453 of transplantation from heart transplant receptors with normal aortic valves (n=2). All patients or 454 close relatives (for autopsy controls) gave written informed consent, and the study was approved

455 by the Health Service Ethics Committees (Comité Ético de Investigación Clínica) of participant 456 centers. The investigation conformed to the principles outlined in the Declaration of Helsinki.

Whole-blood samples. Genes of interest were quantified by quantitative reverse transcription-458 polymerase chain reaction (qRT-PCR) in whole-blood (WB) samples from age-matched case-459 control males (n=104) of the Progression of Early Subclinical Atherosclerosis (PESA). 460 Subclinical cases (PESA-SD) were defined as participants with aortic-valve calcification score 461 >1, as assessed by the Agatston method (1). Controls (PESA-CTL) had no calcification of the 462 aortic valves or coronary arteries. Patients with aortic valve stenosis and moderate to severe 463 calcification (WB-CAVD; n=12) were recruited from Monteprincipe Hospital (Supplementary 464 Table S4, see https://figshare.com/s/4357d56025cb3380a0df). For validation, an independent 465 cohort of mainly severe aortic valve stenosis patients (n=16) was obtained from the MRVALVE 466 study. Whole blood was collected in PAXGene tubes (BDBiosciences) and stored at -80°C 467 before processing.

468 **RNA** isolation

457

469 For each valve, total RNA from a single aortic valve cusp was extracted using the TRIzol LS

470 (Invitrogen, Carlsbad, CA, USA) method and digested with RNase free DNase I (Invitrogen).

471 For PESA and CAVD validation total RNA was using either the PAXgene blood mRNA kit

472 (PreAnalytiX, Hombrechtikon, Switzerland) for manual isolation or the QIAsymphony PAXgene

473 blood RNA kit (PreAnalytiX) for automated isolation using a QIAsymphony SP liquid handling

474 robot (Qiagen, Venlo, Netherlands). RNA purity, concentration, and integrity were assessed by

475 Nanodrop 2000 spectrophotometry (Thermo Scientific) and automated electrophoresis in a 2100

476 Bioanalyzer (RNA6000 Nano LabChip; Agilent). Samples with a RIN > 6 were selected for

477 reverse transcription.

#### 478 **RNA-seq data generation**

479 RNA-seq was performed on DNase I-treated RNA samples with a RIN > 7. RNA-seq libraries 480 were created using the TruSeq Stranded Total RNA with Ribo-Zero Gold Prep Kit (Fetal 481 samples) or TruSeq<sup>TM</sup> RNA Sample Prep Kit v2 (Patient samples; Illumina, Diego, CA). Final 482 cDNA libraries were checked for quality and quantified in the 2100 Bioanalyzer (Agilent). The 483 libraries were loaded in the flow cell at 8 pM, and clusters were generated in Cbot and sequenced 484 as 50bp single-end reads on a Hiseq 2500 instrument (adult samples) or 75 base single-end reads 485 on a Genome Analyzer IIx instrument (fetal samples; Illumina, Diego, CA). Image analysis and 486 base calling were performed using RTA and CASAVA.

#### 487 **Bioinformatics analysis of RNA-Seq.**

#### 488 *Data preprocessing*

489 FASTQ files were pre-processed with cutadapt (16) to eliminate Illumina adapter remains and

- 490 reads were aligned to the human reference transcriptome (GRCh38.78) using RSEM 1.2.31 (14).
- 491 Raw data from 27 RNA-seq samples from GSE76718 were downloaded from GEO and pre-
- 492 processed and aligned using the same pipeline. Control samples were present in all experiments,
- 493 allowing for batch correction of the data using ComBat (12). Data were normalized using TMM
- 494 with log-ratio trimming=0.3 and sumTrim=0.05 and voom-transformed. Only genes with at least
- 495 1 cpm in at least 10 samples were considered for further analysis.

### 496 *Differential Expression Analysis*

- 497 Differential expression between groups was performed using mixed models with eBayes, as
- 498 implemented in the Bioconductor limma package (28). A gene was classified as differentially
- 499 expressed if its Benjamini-Hochberg adjusted P-value was < 0.05 in any of the contrasts under
- 500 study (Supplementary Table S2, see <u>https://figshare.com/s/4357d56025cb3380a0df</u>).
- 501 <u>Clustering and functional analyses</u>
- 502 To identify groups of genes with similar expression profiles across the six conditions under
- 503 consideration, k-means clustering was performed with the R ComplexHeatmap package on all
- 504 expressed genes (19,962 genes). The Elbow method was used to estimate the optimal number of
- 505 <u>clusters, which was set as k=50. The number of genes per cluster was between 215 and 680.</u>
- 506 Ingenuity Pathway Analysis (IPA, Qiagen) was then used to identify significant associations

- 507 between canonical pathways and clusters. Twenty-eight clusters, out of the original collection of
- 508 50, were found to be associated to at least one canonical pathway with Benjamini-Hochberg (B-
- 509 H) adjusted P-value < 0.05 (Supplementary Table S3, see
- 510 https://figshare.com/s/4357d56025cb3380a0df). Concordance between canonical pathway
- 511 profiles was then used to identify groups of clusters that could be associated to similar functions.
- 512 One cluster (cluster 10) did not share any enriched pathway with any other cluster and was not
- 513 considered further. The remaining 27 clusters were grouped into six meta-clusters, using the
- 514 fraction of shared pathways as distance measure. Meta-clusters contained from 332 to 3304
- 515 genes (Figure 2). IPA was used again to re-annotate meta-clusters (Figure 3).

#### 516 Data availability

- 517 Data are deposited in the NCBI GEO database under accession number GSE148219.
- 518 **qRT-PCR for validation studies in whole blood**
- 519 Selection of candidate genes for qRT-PCR. We used whole blood transcriptomics data generated
- 520 in the PESA study to identify genes that might potentially serve as predictors of CAVD. The
- 521 criteria to select these candidate genes were: (i) Genes differentially expressed between
- 522 individuals with (n=59) and without (n=396) calcification of the aortic valve (p-value<=0.05) but
- 523 not differentially expressed between individuals with and without calcification in the coronaries
- 524 (p-value>0.05; Supplementary Table S5, see https://figshare.com/s/4357d56025cb3380a0df).
- 525 (ii) Aortic valve- specific differentially expressed genes with a top 10% correlation between their
- 526 expression level and the level of calcification of the aortic valves. (iii) Their expression levels in
- 527 the RNA-seq samples of the aortic valves were either up-regulated in disease v.s. no disease
- 528 samples or they were upregulated in early embryonic valve samples and in the diseased samples
- 529 but not in the control samples. From the 10 genes fulfilling the three criteria, 6 were selected
- 530 based on their presumptive function (Supplementary Table S6, see
- 531 <u>https://figshare.com/s/4357d56025cb3380a0df).</u>

532 qRT-PCR. RNA samples from PESA and CAVD patients (75ng) was reverse transcribed using 533 the High Capacity cDNA Reverse Transcription kit with RNase Inhibitor (ThermoFisher 534 Scientific). Gene expression was quantified by qPCR using TaqMan probes (ThermoFisher 535 Scientific; Supplementary Table S8) in 52 case-control pairs from the PESA study (case: positive 536 calcification of the aortic valve) and in 12 CAVD patients from Monteprincipe Hospital. Each 537 reaction was performed in triplicate and the final reaction volume was set to 10µl. qPCRs were 538 conducted in 384-well plates (Cat. no. 4344345; Applied Biosystems) using an ABI PRISM 539 7900HT Sequence Detection System (Applied Biosystems). Each well included 3µl (2.25 ng) 540 cDNA, 5ul 2x TaqMan Universal Mater Mix (ThermoFisher Scientific), 0.5ul 20x TaqMan 541 assay (ThermoFisher Scientific), and 1.5µl H<sub>2</sub>O. The cycling protocol was 10 min at 95°C 542 followed by 60 cycles of 15 s at 95°C and 60 s at 60°C. Baseline and threshold were set 543 automatically in SD S 2.3 (ThermoFisher Scientific) to calculate Cq values. The Cq values were 544 then imported into qBasePlus (Biogazelle, Zwijnaarde, Belgium) to calculate relative quantities 545 normalised to GAPDH. Data were scaled per sample to eliminate potential biases.

#### 546 Statistical analysis

547 A linear mixed model was used to assess correlation between the expression of each gene in WB 548 and the degree of calcification in the aortic valve for each disease group (PESA-SD, WB-549 CAVD) separately. Visualization was performed using the multiplot function and the ggplot2 R 550 package. All samples were jointly analyzed using a linear mixed model with the origin of the 551 samples as random effect. The same analysis was performed to assess the correlation of the gene 552 expression with calcification of the coronaries. Four genes which showed no correlation with the 553 amount of calcium in the coronaries were selected to build a gene expression signature using a 554 multinomial model with interaction for the prediction of the three groups of individuals with 555 different disease severity WB-CAVD>PESA-SD>PESA-CTL. To quantify the predictive power

of the gene signature between different groups of individuals a logistic regression model was fit
to the data and the corresponding ROC curves were built using the ROCR R package.

558

559

561

### 560 Acknowledgments

562 We thank S. Bartlett for English editing. This work was supported by the Ministerio de Ciencia, 563 Innovación y Universidades from Spain, (grants SAF2016-78370-R, CB16/11/00399, CIBER 564 CV, and RD16/0011/0021, TERCEL) and the Fundación BBVA (grant BIO14 298) and 565 Fundación La Marató TV3 (grant 20153431) to J.L.dlP; and the Fondation pour la Recherche 566 Médicale, grant from l'Institut National de la Santé et de la Recherche Médicale (grant 567 DPC20111123002) and l'Association Française contre les Myopathies (grant TRIM-RD) to 568 S.Z. The cost of this publication was supported in part with funds from the ERDF. The CNIC is 569 supported by the Ministerio de Ciencia, Innovación y Universidades (MCNU), the Instituto de 570 Salud Carlos III (ISCIII), and the Pro CNIC Foundation, and is a Severo Ochoa Center of 571 Excellence (grant SEV-2015-0505).

#### 572 Author contributions

573 Donal MacGrogan, PhD: Design of work: Substantial; interpretation of data: Substantial; 574 Writing: original draft: Lead; Writing: review & editing: Lead. final approval of the version to be

575 published.

576 Beatriz Martínez Poveda, PhD: Investigation: Lead; Methodology: Substantial; final approval of 577 the version to be published.

578 Jean-Pierre Desvignes, MD: Investigation: Supporting; final approval of the version to be 579 published.

580 Leticia Fernandez-Friera, MD, PhD: Data curation: Substantial; Formal analysis: Substantial;

581 Investigation: Supporting; final approval of the version to be published.

582 Manuel José Gomez, PhD: Formal analysis: Substantial; Visualization: Substantial; final

583 approval of the version to be published.

- 584 Eduardo Gil Vilariño, BSc: Investigation: Supporting; final approval of the version to be 585 published.
- 586 Sergio Callejas Alejano, BSc: Investigation: Supporting; final approval of the version to be 587 published.
- Pablo Garcia-Pavia, MD, PhD: Investigation: Supporting; final approval of the version to bepublished.
- 590 Jorge Solis, MD, PhD: Investigation: Supporting; final approval of the version to be published.
- 591 Joaquín Lucena, MD, PhD:Investigation: Supporting; final approval of the version to be 592 published.
- 593 David Salgado, PhD: Data curation: supporting; Formal analysis: supporting; Investigation:
- supporting; final approval of the version to be published.

595 Gwenaelle Collod-Béroud, PhD: Formal analysis: Supporting; Investigation: Supporting; final 596 approval of the version to be published.

- 597 Emilie Faure, PhD: Formal analysis: Supporting; Investigation: Supporting; final approval of the
- 598 version to be published.
- 599 Alexis Théron, MD: Investigation: Supporting; final approval of the version to be published.
- 600 Julia Torrents, MD: Investigation: Supporting; final approval of the version to be published.
- 601 Jean-François Avierinos, MD: Investigation: Supporting; final approval of the version to be 602 published.
- 603 Lorena Montes, MD: Investigation: Supporting; final approval of the version to be published.
- 604 Ana Dopazo, PhD: Investigation: Supporting; Supervision: Supporting; final approval of the 605 version to be published.
- 606 Valentín Fuster, MD, PhD: Investigation: Supporting; Resources: Supporting; final approval of
- 607 the version to be published.
- 608 Borja Ibañez, MD, PhD: Investigation: Supporting; Writing review & editing: Supporting;
- 609 final approval of the version to be published.
- 610 Fátima Sánchez-Cabo, PhD: Formal analysis: Equal; Writing review & editing: Supporting;
- 611 final approval of the version to be published.
- 612 Stephane Zaffran, PhD: Investigation: Supporting; Supervision: Supporting; Writing –review &
- 613 editing: Supporting; final approval of the version to be published.
- 614 José Luis de la Pompa, PhD: Conceptualization: Lead; Funding acquisition: Lead; Project
- 615 administration: Lead; Resources: Lead; Supervision: Equal; Visualization: Equal; Writing -
- 616 review & editing: Supporting.

<sup>4</sup> 

## **Competing interests**

619 The authors declare no competing interests.

#### 623 **REFERENCES**

Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Jr., and Detrano 624 1. 625 **R**. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll 626 Cardiol 15: 827-832, 1990. 627 Beppu S. Suzuki S. Matsuda H. Ohmori F. Nagata S. and Miyatake K. Rapidity of 2. 628 progression of aortic stenosis in patients with congenital bicuspid aortic valves. Am J Cardiol 71: 629 322-327, 1993. 630 Chang HC, and Guarente L. SIRT1 and other sirtuins in metabolism. Trends 3. 631 Endocrinol Metab 25: 138-145, 2014. 632 Coffey S, Cairns BJ, and Iung B. The modern epidemiology of heart valve disease. 4. 633 Heart 102: 75-85, 2016. 634 Fernandez-Ortiz A, Jimenez-Borreguero LJ, Penalvo JL, Ordovas JM, Mocoroa A, 5. Fernandez-Friera L, Laclaustra M, Garcia L, Molina J, Mendiguren JM, Lopez-Melgar B, 635 636 de Vega VM, Alonso-Farto JC, Guallar E, Sillesen H, Rudd JH, Fayad ZA, Ibanez B, Sanz 637 G, and Fuster V. The Progression and Early detection of Subclinical Atherosclerosis (PESA) 638 study: rationale and design. Am Heart J 166: 990-998, 2013. 639 Freeman RV, and Otto CM. Spectrum of calcific aortic valve disease: pathogenesis, 6. 640 disease progression, and treatment strategies. Circulation 111: 3316-3326, 2005. 641 Gottlieb Sen D, Halu A, Razzaque A, Gorham JM, Hartnett J, Seidman JG, Aikawa 7. 642 E, and Seidman CE. The Transcriptional Signature of Growth in Human Fetal Aortic Valve Development. Ann Thorac Surg 106: 1834-1840, 2018. 643 644 Guauque-Olarte S, Droit A, Tremblay-Marchand J, Gaudreault N, Kalavrouziotis 8. 645 D, Dagenais F, Seidman JG, Body SC, Pibarot P, Mathieu P, and Bosse Y. RNA expression 646 profile of calcified bicuspid, tricuspid, and normal human aortic valves by RNA sequencing. 647 *Physiol Genomics* 48: 749-761, 2016. 648 Hetz C. The unfolded protein response: controlling cell fate decisions under ER stress 9. 649 and beyond. Nat Rev Mol Cell Biol 13: 89-102, 2012. 650 10. Hinton RB, and Yutzey KE. Heart valve structure and function in development and 651 disease. Annu Rev Physiol 73: 29-46, 2011. 652 Hulin A, Hego A, Lancellotti P, and Oury C. Advances in Pathophysiology of Calcific 11. 653 Aortic Valve Disease Propose Novel Molecular Therapeutic Targets. Front Cardiovasc Med 5: 654 21, 2018. 655 12. Johnson WE, Li C, and Rabinovic A. Adjusting batch effects in microarray expression 656 data using empirical Bayes methods. Biostatistics 8: 118-127, 2007. 657 13. Krishnan A, Samtani R, Dhanantwari P, Lee E, Yamada S, Shiota K, Donofrio MT, 658 Leatherbury L, and Lo CW. A detailed comparison of mouse and human cardiac development. 659 Pediatr Res 76: 500-507, 2014. 660 Li B, and Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data 14. 661 with or without a reference genome. BMC bioinformatics 12: 323, 2011. 662 Lincoln J, Lange AW, and Yutzey KE. Hearts and bones: shared regulatory 15. 663 mechanisms in heart valve, cartilage, tendon, and bone development. Dev Biol 294: 292-302, 664 2006. 665 Martin M. CUTADAPT removes adapter sequences from high-throughput sequencing 16. 666 reads. EMBnetjournal 17: 10-12, 2011. 667 17. Mathieu P, Bouchareb R, and Boulanger MC. Innate and Adaptive Immunity in 668 Calcific Aortic Valve Disease. J Immunol Res 2015: 851945, 2015. 669 Mohler ER, 3rd, Gannon F, Reynolds C, Zimmerman R, Keane MG, and Kaplan 18.

670 **FS**. Bone formation and inflammation in cardiac valves. *Circulation* 103: 1522-1528, 2001.

671 19. Morita M, Gravel SP, Hulea L, Larsson O, Pollak M, St-Pierre J, and Topisirovic I. 672 mTOR coordinates protein synthesis, mitochondrial activity and proliferation. Cell Cycle 14: 673 473-480, 2015. 674 20. Niehrs C, and Pollet N. Synexpression groups in eukaryotes. *Nature* 402: 483-487, 1999. 675 676 21. Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP, 3rd, Guvton RA, O'Gara PT, Ruiz CE, Skubas NJ, Sorajja P, Sundt TM, 3rd, Thomas JD, Anderson JL, 677 678 Halperin JL, Albert NM, Bozkurt B, Brindis RG, Creager MA, Curtis LH, DeMets D, 679 Guyton RA, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, 680 Stevenson WG, Yancy CW, American College of C, American College of 681 Cardiology/American Heart A. and American Heart A. 2014 AHA/ACC guideline for the 682 management of patients with valvular heart disease: a report of the American College of 683 Cardiology/American Heart Association Task Force on Practice Guidelines. J Thorac 684 Cardiovasc Surg 148: e1-e132, 2014. 685 22. **O'Brien KD**. Pathogenesis of calcific aortic valve disease: a disease process comes of 686 age (and a good deal more). Arterioscler Thromb Vasc Biol 26: 1721-1728, 2006. 687 23. Olsson M, Dalsgaard CJ, Haegerstrand A, Rosenqvist M, Ryden L, and Nilsson J. 688 Accumulation of T lymphocytes and expression of interleukin-2 receptors in nonrheumatic 689 stenotic aortic valves. J Am Coll Cardiol 23: 1162-1170, 1994. 690 Otto CM, Kuusisto J, Reichenbach DD, Gown AM, and O'Brien KD. 24. 691 Characterization of the early lesion of 'degenerative' valvular aortic stenosis. Histological and 692 immunohistochemical studies. Circulation 90: 844-853, 1994. 693 25. Otto CM, and Prendergast B. Aortic-valve stenosis--from patients at risk to severe 694 valve obstruction. N Engl J Med 371: 744-756, 2014. 26. 695 Padang R, Bagnall RD, Tsoutsman T, Bannon PG, and Semsarian C. Comparative 696 transcriptome profiling in human bicuspid aortic valve disease using RNA sequencing. Physiol 697 Genomics 47: 75-87, 2015. 698 27. Pearce EL. Metabolism in T cell activation and differentiation. Curr Opin Immunol 22: 314-320, 2010. 699 700 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK. limma 28. 701 powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic 702 acids research 43: e47, 2015. 703 29. Roberts WC, and Ko JM. Frequency by decades of unicuspid, bicuspid, and tricuspid 704 aortic valves in adults having isolated aortic valve replacement for aortic stenosis, with or 705 without associated aortic regurgitation. Circulation 111: 920-925, 2005. 706 30. Schlotter F, Halu A, Goto S, Blaser MC, Body SC, Lee LH, Higashi H, DeLaughter 707 DM, Hutcheson JD, Vyas P, Pham T, Rogers MA, Sharma A, Seidman CE, Loscalzo J, 708 Seidman JG, Aikawa M, Singh SA, and Aikawa E. Spatiotemporal Multi-Omics Mapping 709 Generates a Molecular Atlas of the Aortic Valve and Reveals Networks Driving Disease. 710 Circulation 138: 377-393, 2018. 711 31. Siu SC, and Silversides CK. Bicuspid aortic valve disease. J Am Coll Cardiol 55: 2789-712 2800, 2010. 713 Small A, Kiss D, Giri J, Anwaruddin S, Siddiqi H, Guerraty M, Chirinos JA, 32. 714 Ferrari G, and Rader DJ. Biomarkers of Calcific Aortic Valve Disease. Arterioscler Thromb 715 Vasc Biol 37: 623-632, 2017. 716 33. Steiner I, Krbal L, Rozkos T, Harrer J, and Laco J. Calcific aortic valve stenosis: 717 Immunohistochemical analysis of inflammatory infiltrate. Pathol Res Pract 208: 231-234, 2012. 718 Winchester R, Wiesendanger M, O'Brien W, Zhang HZ, Maurer MS, Gillam LD, 34. 719 Schwartz A, Marboe C, and Stewart AS. Circulating activated and effector memory T cells are 720 associated with calcification and clonal expansions in bicuspid and tricuspid valves of calcific 721 aortic stenosis. J Immunol 187: 1006-1014, 2011.

- 722 35. Wu HD, Maurer MS, Friedman RA, Marboe CC, Ruiz-Vazquez EM,
- 723 Ramakrishnan R, Schwartz A, Tilson MD, Stewart AS, and Winchester R. The lymphocytic
- infiltration in calcific aortic stenosis predominantly consists of clonally expanded T cells. J
   *Immunol* 178: 5329-5339, 2007.
- 726 36. Zhao Y, Nicoll R, He YH, and Henein MY. The effect of statins on valve function and
- 727 calcification in aortic stenosis: A meta-analysis. *Atherosclerosis* 246: 318-324, 2016.

### 729 FIGURE LEGENDS

731	Figure 1. Aortic valve collection summary, bioinformatics pipeline and mapping of sample
732	batches and biological conditions in gene expression profile-based principal component analysis
733	(PCA) plots. (A) Study outline summarizing aortic valve collection. During the study period,
734	aortic valve leaflets were collected from human fetuses at gestational week (week) 9 (n=4), 12-
735	13 (n=3), and 22 (n=2), adult controls (n=8) and patients with calcified BAV (cBAV; n=5) or
736	TAV (cTAV; n=7). Samples were processed and subject to RNA-seq to characterize their gene
737	expression profiles. Complementary expression data from control, cBAV and cTAV samples
738	was downloaded from GEO (GSE76718). The resulting data set collection was analyzed with a
739	bioinformatics pipeline that included preprocessing of fastq files with cutadapt, alignment
740	against a human transcriptome reference and estimation of gene expression levels with RSEM,
741	and differential expression testing with limma, considering that samples could be ascribed to five
742	experimental batches. Expression profiles were clustered using k-means to obtain a collection of
743	50 clusters, which were then annotated with the Canonical Pathway feature of IPA. Twenty-
744	seven clusters associated to at least one canonical pathway (with Benjamini-Hochberg adjusted
745	p-value $< 0.05$ ) were grouped into six meta-clusters on the basis of concordant annotations.
746	Meta-clusters were finally re-annotated with IPA. (B) Hierarchical clustering of the samples
747	overall gene expression based on Euclidean Distance (C) Gene expression profile PCA plot
748	mapping of biological conditions. Sample points are colored according to fetal developmental
749	stage or adult disease state (no disease, n=16; cBAV, n=15 and cTAV, n=16). In PC1, control
750	samples separate from fetal samples and cBAV or cTAV samples. In PC2, early-stage fetal
751	samples (weeks 9 and 13) separate from adult control and disease samples. Percentages on axis
752	labels denote the amount of variance explained by each principal component.

754	Figure 2. Concordance analysis of k-means clustering. K-means-based gene clusters (see
755	Supplementary Fig. S 1B, see https://figshare.com/s/d9a68e0d6faa67e4c6c3) were grouped into
756	metaclusters according to the agreement of their enriched canonical pathways. The white to
757	violet squares in the central heatmap represent the percentage (from 0 to 100) of Canonical
758	Pathways (CP) shared between each pair of 50 k-means clusters for which at least one enriched
759	and shared term was found (k=27 out of the original 50 clusters). Hierarchical clustering of the
760	concordance matrix identified six meta-clusters (MC1-6) The heatmap on the left represents the
761	mean gene expression values for each cluster across biological conditions, on a normalized scale
762	(from -2 to 2, for values below and above the mean, respectively). The dendrogram at the top
763	represents functional relatedness between clusters, which was used to group them into meta-
764	clusters. Meta-clusters are identified with colored boxes (MC1 to MC6, in red, blue, green,
765	purple, orange and yellow, respectively) and represent groups of clusters associated to similar
766	functions.

- Figure 3. Functional analysis of meta-clusters. The six gene metaclusters (MC1-6) described in
   Figure 2 were annotated with IPA. Barplots on the right-hand side represent enrichment values
- 770 for the top five IPA Canonical Pathways associated with each meta-cluster (MC), selected
- 771 according to enrichment significance and having an adjusted Benjamini-Hochberg P<0.05. Meta-
- 772 clusters are identified with colored boxes (MC1 to MC6, in red, blue, green, purple, orange and
- 773 yellow, respectively) and represent groups of clusters associated to similar functions. Central
- 774 boxes list all the genes associated with the selection of pathways associated with each
- 775 metacluster. The heatmap on the left represents normalized expression values for individual
- genes across biological conditions, on a normalized scale (from -2 to 2, for values below and
- 777 <u>above the mean, respectively).</u>
- 778

779	Figure 4. A gene expression signature predictive of CAVD. (A) Correlation of qPCR data (y-
780	axis) vs aortic-valve calcification levels (x-axis) for the 6 selected genes in whole blood (WB).
781	Each point represents an individual's level of gene expression (y-axis) and its aortic valve
782	calcification (x-axis). Dots are colored based on the group to which the individual belongs:
783	PESA-CTL (n=52), PESA-SD (n=52) or WB-CAVD (n=12). Solid lines represent the linear
784	model fit for PESA-SD and WB-CAVD patients. Shaded regions represent the confidence of the
785	linear regression estimate for each calcium level. (B) Receiver operating characteristic (ROC)
786	curve of multivariate logistic regression model predicting CAVD (WB-CAVD) vs preclinical
787	aortic-valve calcification (PESA-SD; thick discontinuous line, AUC=0.97) vs valve calcification
788	(PESA-CTL; thick solid line, AUC=0.91). A third ROC curve for the prediction of valve
789	calcification (PESA-SD) vs no calcification (PESA-CTL; thin solid line, AUC=0.71) is shown.
790	The discontinuous diagonal line indicates a fifty-fifty percent chance to predict each group. (C)
791	Follow-up of the model in PESA subgroups after 3 years. Validation in an independent group of
792	CAVD patients (WB-CAVD, n=16) vs a PESA-SD subgroup (n=25; thick discontinuous line,
793	AUC=0.90) vs a PESA-CTL subgroup (n=35; thick solid line, AUC=0.88). A third ROC curve
794	for the prediction of PESA-SD (n=25) vs PESA-CTL (n=35; thin solid line, AUC=0.71) is
795	shown. The thin discontinuous diagonal line indicates a fifty-fifty percent chance to predict each
796	group.





Figure\_1\_MacGrogan\_et\_al



Figure\_2\_MacGrogan\_et\_al



Figure\_3\_MacGrogan\_et\_al







Figure 4\_MacGrogan et al

Α