

ARTICLE OPEN



Mapping drug-target interactions and synergy in multi-molecular therapeutics for pressure-overload cardiac hypertrophy

Aparna Rai¹, Vikas Kumar², Gaurav Jerath^{1,3}, C. C. Kartha⁴✉ and Vibin Ramakrishnan¹✉

Advancements in systems biology have resulted in the development of network pharmacology, leading to a paradigm shift from “one-target, one-drug” to “target-network, multi-component therapeutics”. We employ a chimeric approach involving in-vivo assays, gene expression analysis, cheminformatics, and network biology to deduce the regulatory actions of a multi-constituent Ayurvedic concoction, Amalaki Rasayana (AR) in animal models for its effect in pressure-overload cardiac hypertrophy. The proteomics analysis of in-vivo assays for Aorta Constricted and Biologically Aged rat models identify proteins expressed under each condition. Network analysis mapping protein–protein interactions and synergistic actions of AR using multi-component networks reveal drug targets such as ACADM, COX4I1, COX6B1, HBB, MYH14, and SLC25A4, as potential pharmacological co-targets for cardiac hypertrophy. Further, five out of eighteen AR constituents potentially target these proteins. We propose a distinct prospective strategy for the discovery of network pharmacological therapies and repositioning of existing drug molecules for treating pressure-overload cardiac hypertrophy.

npj Systems Biology and Applications (2021)7:11; <https://doi.org/10.1038/s41540-021-00171-z>

INTRODUCTION

Modern medicine is primarily driven by the discovery of small-molecule entities with pharmacological actions¹. Despite the vast scope of the chemical universe, the probability of success in discovering a new molecule with potential therapeutic effects is now becoming more and more challenging. In addition, the “one drug, one target” mode of drug action cannot generally lead to multiple effects in complex or multifactorial diseases because of the underlying complexity of biological networks^{2–5}. The disease is known to be a rare consequence of an abnormality in a single gene or gene product but reflects the perturbations of a complex intracellular and intercellular network in organ systems. The emerging tools of network medicine offer a platform to explore not only the molecular complexity of a particular disease, leading to the identification of disease modules and pathways, but also in investigating molecular defects among apparently distinct pathological phenotypes^{6–8}. A sensible approach for the treatment of complex diseases is to have a combinatorial drug with constituents that target multiple pathways in a disease-specific network^{9,10}.

The complexity of the underlying biological processes and interactome, therefore, demands multiple synergistic therapeutics targeting different proteins involved in the disease onset and prognosis. Identification of such synergistic therapeutic partners is both complex and arduous due to various drug–drug interactions. We believe that it would be prudent to identify traditional medicines in eastern methods, which have been practiced over centuries, for their potential utility¹¹. To bring such therapeutic solutions to the platform, modern medicine can be a rewarding exercise. As a test case, we have identified Amalaki Rasayana (AR), a concoction used for the treatment of cardiovascular diseases, diabetes, and rejuvenation therapies in Ayurveda, with an

objective to re-invent its efficacy through established analytical procedures of modern medicine.

Cardiac hypertrophy due to pressure overload or pressure-overload left ventricular cardiac hypertrophy (LVCH), is one of the leading causes of mortality in recent times, especially in the younger population and patients having hypertension^{12,13}. The physiopathological characterization of hypertrophy is multifactorial and it involves a higher degree of complexity at the cellular and molecular level, across multiple signaling pathways¹⁴. Pressure-overload LVCH is a significant phenomenon in cardiac hypertrophy and accounts for 10–15% cases in the adult population and ~40% with hypertension, leading to adverse cardiovascular events such as heart failure and sudden death¹⁵. Therapeutic interventions such as the use of ACE and myosin inhibitors and various anti-hypertensive drugs have significant clinical benefits but have limited success due to the ceiling effect in many patients^{16,17}. Therefore, extensive research on the molecular basis of this complex, multifaceted physiopathological malady at the levels of pathways, drug dosage, and targeted therapy can be helpful in proposing new therapeutic options^{18,19}.

Interestingly, Ayurveda, the ancient Indian system of medicine employs combinatorial therapy^{11,20}. Ayurvedic medicines have been refined and evolved continuously over centuries and is one of the most dependent forms of alternative medicine for more than 1.5 billion population living in the Indian sub-continent. Ayurvedic medicines commonly have multiple components derived from natural sources and are believed to work through a cooperative mechanism or “synergy” to restore the balance of life and body functions through heightened physiological response, not achievable by individual constituents^{21,22}. Synergistic action maybe because of the involvement of a set of targeted proteins, related mechanisms, or drug combinations. In modern

¹Molecular Informatics and Design Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati, Assam, India. ²Laboratory of Cardiovascular Science, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA. ³PepThera Laboratories Private Limited, Guwahati, Assam, India. ⁴Society for Continuing Medical Education & Research, Kerala Institute of Medical Sciences, Thiruvananthapuram, Kerala, India. ✉email: cckartha@gmail.com; vibin@iitg.ac.in

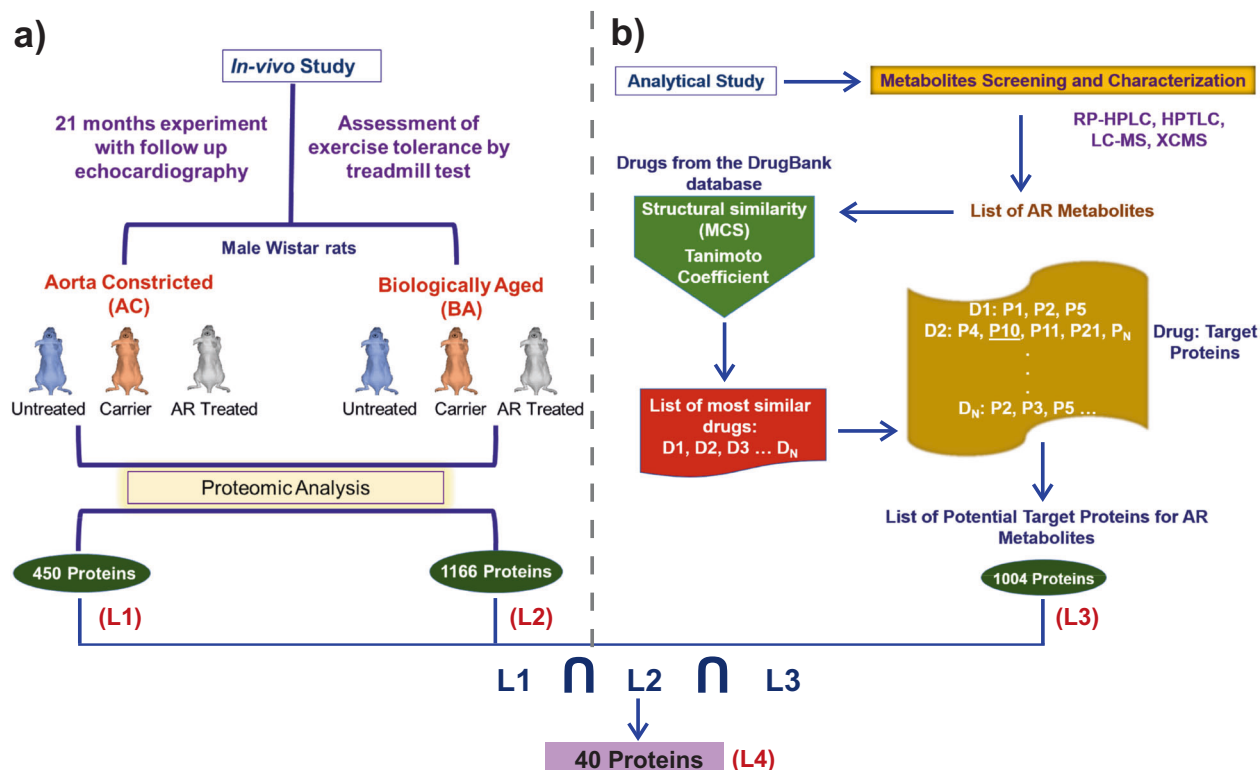


Fig. 1 in-vivo studies and data extraction. **a** Illustration presenting an overall schematic representation of the in-vivo study performed on male Wistar rats for 21 months²⁸. The proteins from the AC and BA samples were identified by proteomic analysis and designated as L1 and L2, respectively. **b** The AR composition was identified by mass spectroscopy techniques. The AR constituent metabolites, thus identified were used to search structurally similar drugs in the DrugBank database. Further, the targets corresponding to the highest-ranked drug molecules were identified to relate their activity against each metabolite (L3). The intersection of L1, L2, and L3 had 40 proteins (L4). These proteins represent the set of proteins targeted by various AR metabolites, which leads to the drug action of the AR concoction. Therefore, we used list L4 for our further analysis.

drugs, synergistic action is due to their combinatorial effects on multiple targets identified by their network topological features. Synergistic action predictions are generally pathway-based, drug similarity-based, and using omics-based models^{23–26}. Advancement in systems biology has resulted in a new concept known as “network pharmacology”, which offers a radical shift from the existing paradigm of “one-target, one-drug” mode to a “network-target, multiple-component-therapeutics”²⁷. We present a unique approach, integrating the tools of cheminformatics, proteomics, in-vivo experiments, and network analysis to identify biological targets and synergy of constituents in a typical multi-molecular medicinal formulation for complex diseases. We analyzed the cumulative in-vivo effect of AR, a concoction used by practitioners of Ayurveda for centuries for the treatment of pressure-overload LVCH²⁸. The data generated through in-vivo studies were utilized to detail the possible drug targets for each of the constituent metabolites of AR. We also identified those AR metabolites, structurally similar to known drugs and potential targets in modern medicine. Further, we also identified a set of existing drug molecules that could be re-purposed for pressure-overload LVCH.

RESULTS

in-vivo studies for AR efficacy

AR is a result of a series of systematic preparatory procedures, prepared primarily from the fruits of *Phyllanthus Emblica* or *Emblica Officinalis* (as major constituent) as per the guidelines of Charaka Samhita, written in the third century BCE²⁹. In an earlier experiment, spanning a total period of 21 months, we considered

two groups of Wistar rats; (i) Aorta Constricted (AC) with pressure-overload LVCH, induced by clipping ascending Aorta with titanium clips, and (ii) Biologically Aged (BA) rats (Fig. 1a). Both these groups individually were further sub-grouped, where these animals were either given AR (AR treated, Test sub-group) or ghee and honey treated (carrier sub-group) orally by gavage regularly. Another sub-group in both the AC and BA groups acted as control (Untreated sub-group), where no drug was administered to the animals²⁸. Thereafter, we evaluated the effects of long-term administration of AR on the structure and function of the heart. Histology, as well as gene and protein expression analysis, were done in left ventricular heart tissues, collected after the sacrifice of the animals. In the 21-months experiment on Wistar rats, we found that AR intake resulted in an improved left ventricular function and decreased left ventricular hypertrophy in AC rats. AC groups fed AR also had increased fatigue time in the treadmill exercise test. BA group fed AR also had improved left ventricular function. Further, the protein expression profile of the heart tissue of both AR and BA groups with AR intake revealed upregulation of SERCA2, CaM, Myh11 as well as antioxidant, autophagy, oxidative phosphorylation, and tricarboxylic acid (TCA) cycle proteins. ADRB1/2 and pCREB expressions were also increased, but the expressions of pAMPK and NF-κB were decreased. The protein expression analysis indicates that AR has a beneficial effect on myocardial energetics, muscle contractile function, and exercise tolerance capacity in rats²⁸. We identified 450 proteins (List L1) in cardiac tissues of AC rats, and 1166 proteins in cardiac tissues of BA rats (List L2). The Lists L1 and L2 along with their expression profiles are given in the Supplementary Dataset S1.

Table 1. AR constituent metabolites.

Serial no.	Metabolite	Associated pathways/functions
1.	Gallic acid	Mucous protection, astringent effects
2.	Ellagic acid	Antioxidant and anti-proliferative effects
3.	Biocytin	Vitamin H (biotin) metabolism
4.	Methylcobalamin	Vitamin B12 (cyanocobalamin) metabolism
5.	L-methionine	Amino acid metabolism
6.	Pyridoxamine phosphate	Vitamin B6 (pyridoxine) metabolism
7.	Prostaglandin B1	Prostaglandin formation from arachidonate
8.	4-Hydroxy-all-trans-retinyl_acetate	Vitamin A (retinol) metabolism
9.	13'-carboxy-alpha-tocotrienol	Vitamin E metabolism
10.	Cholic acid	Bile acid metabolism
11.	1alpha,24 R,25-trihydroxyvitamin D3 (calcitriol)	Vitamin D3 (cholecalciferol) metabolism
12.	Guanylic acid	Purine metabolism
13.	Carbamoyl phosphate	Pyrimidine metabolism
14.	5-formiminotetrahydrofolate	Vitamin B9 (folate) metabolism, Histidine metabolism
15.	Nicotinate D-ribonucleoside	Vitamin B3 (nicotinate and nicotinamide) metabolism
16.	Sulfate derivative of norepinephrine (noradrenaline sulfate)	Tyrosine metabolism
17.	12-oxo-20-dihydroxy-leukotriene_B4	Leukotriene metabolism
18.	5S,6S-epoxy-15S-hydroxy-7E,9E,11Z,13E-eicosatetraenoic acid	Arachidonic acid metabolism or putative anti-inflammatory metabolite

The list of AR constituent metabolites identified from HPLC, HPTLC, and LC-MS analysis.

In addition to the above experiment, AR was characterized by a chemical composition analysis using HPLC, HPTLC, and LC-MS which yielded 18 constituent Metabolites (Table 1). These constituents were found to have biological importance in the cardiovascular system. The HPTLC profiles of samples of the finished formulation indicate the presence of gallic acid and ellagic acid (Supplementary Fig. S1). The LC-MS analysis of lyophilized powder of AR revealed the enrichment of components such as putative anti-inflammatory arachidonate (eicosatetraenoic acid), norepinephrine sulfate, and vitamin metabolites, identified using XCMS software used for metabolomics study³⁰ (Supplementary Tables S1 and S2).

Drug-metabolite similarity search

The cheminformatics analysis in this study was firmly grounded on the basic idea that similar structures, with chemically identical functional groups, tend to possess similar functions^{31–33}. Considering the list of AR metabolites identified earlier (Table 1)²⁸, we screened the DrugBank database with 9000 approved and experimental drugs to identify molecules that are structurally similar to AR metabolites³⁴. We compared the two-dimensional structure of drugs with those of AR metabolites using fragment-based drug discovery. Our approach includes the identification of maximum common substructure (MCS) between the metabolites and known drug molecules through local and global structural similarity search using the Tanimoto Coefficient (Tc) (Fig. 1b)³⁵. We used a Tc cutoff of 0.6 for quantifying the structural similarity between two molecules. The screened drug molecules were further verified manually by visualizing the pairwise similarity between drug molecules and the AR metabolites using Small Molecular Subgraph Detector (SMSD) toolkit³⁶. The list of AR metabolites and drugs screened are given in the Supplementary Dataset S1. Following the basic paradigm of a similar structure can lead to similar mechanisms of action; targets of structurally similar drugs can also be the targets for the AR metabolites. We thus obtained a list of AR metabolites and their associated potential target proteins (L3). Altogether, 1004 targets were identified (Supplementary Dataset S1). This list includes a set of targets of

known drugs, which can be potentially targeted by the AR metabolites as well.

Protein-protein interaction network analysis screen important proteins

Network biology has been widely used in understanding the complexity of various diseases, their interacting patterns, role, and the importance of interaction patterns^{37–41}. Lists L1 and L2 represent the proteins, whose expression was perturbed due to the administration of AR. The L3 list described above-provided information on potential drug targets that can also interact with the AR metabolites. Therefore, the intersection of the three lists (L1, L2, and L3) would represent the list of proteins responsible for the disease prognosis, which when targeted through AR metabolites lead to the therapeutic action. The intersection of the three lists (L4) has 40 proteins (Supplementary Dataset S1). Next, we constructed the protein-protein interaction network of the L4 proteins and analyzed the network for various topological properties (Fig. 2a), with an aim to identify the proteins responsible for network/pathway integrity.

The connected component of the PPI network (N) constructed using L4 proteins comprised of 38 nodes (proteins) and 121 edges (interactions). The interaction data of the proteins in the form of an adjacency list can be found in Supplementary Dataset S1. The p-value of the PPI enrichment according to the STRING database was 1.0e-16, which indicated that the interaction data are statistically significant. Such an enrichment score suggested that proteins have more interactions among themselves than what would be expected for a random set of proteins of similar size and are biologically significant as a group. To get a deeper insight, the network was analyzed for various structural/topological properties listed in Fig. 2b and Supplementary Table S3. The fundamental property, average degree, $\langle k \rangle$ of the network was 6, indicating interconnectedness of the nodes in the network. The degree of fourteen nodes higher than the average degree in the network suggests their function in promoting the network's robustness through their connectedness against random external perturbations^{42,43}. This was further confirmed by pathway analysis

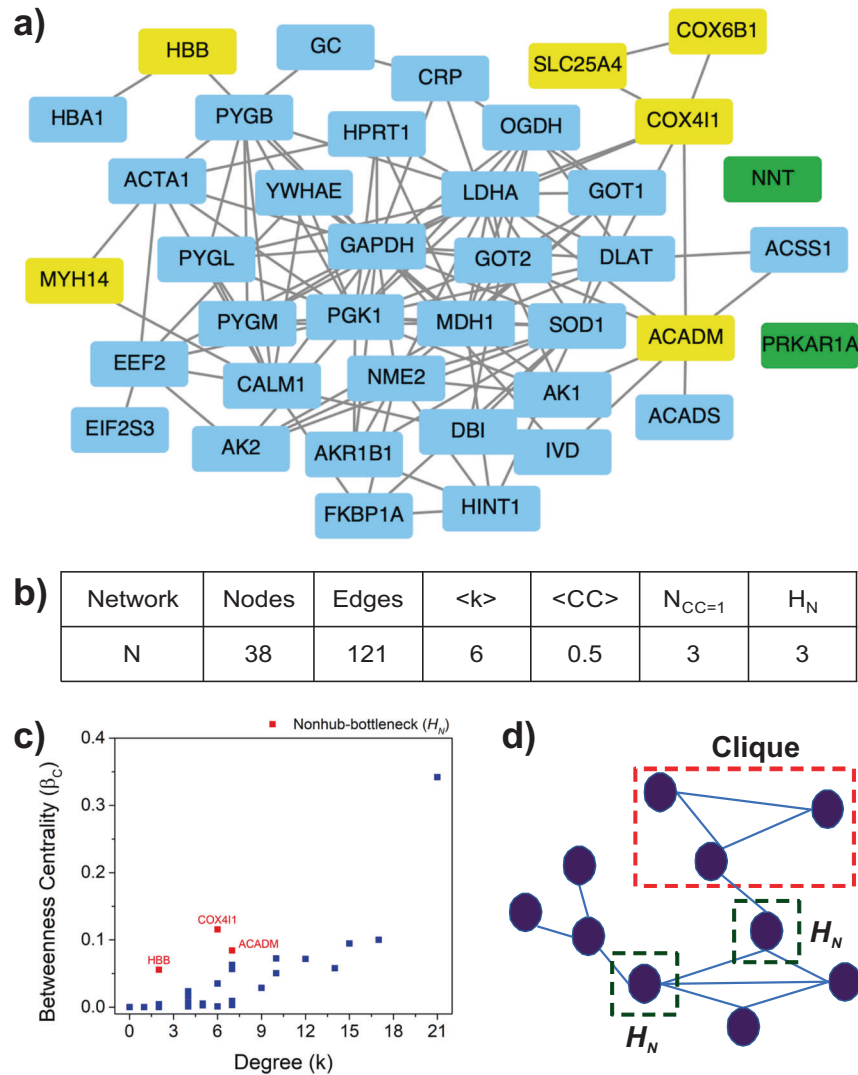


Fig. 2 PPI network. Using STRING database, PPI network for 40 common proteins was constructed. **a** The network has 38 nodes and 121 edges. Nodes namely PRKAR1A and NNT (Green Boxes) do not interact with any of the other nodes. For statistical analysis, the largest connected component was considered and the unconnected nodes were ignored. The yellow boxes are the set of targets obtained from network analysis which were further studied for their therapeutic potential against cardiovascular diseases. **b** The table illustrates various structural properties of the network corresponding to 40 common proteins, where N refers to the name of the network followed by the number of nodes and edges in the network. $\langle k \rangle$ represents the average degree of the network along with the average clustering coefficient ($\langle CC \rangle$). $N_{CC=1}$ is the number of nodes having $CC=1$, whereas H_N represents the nonhub bottlenecks that have a very low degree but high betweenness centrality. **c** The betweenness centrality vs degree curve enlisting nonhub bottlenecks (H_N) as red dots in the graph. **d** Schematic representation of clique structures (red dotted rectangle) and H_N (green dotted square) property in a network. Cliques are closed subgraphs in a network that are vital to the network stability and robustness. H_N represents the nodes that act as bridges between consecutive network components like hub nodes and subgraphs. Targeting such nodes helps to maintain network integrity and, therefore, is hypothesized to be ideal pharmacological drug targets in the network.

using the STRING database, where the cutoff for false discovery rate (FDR) was <0.05 (95% significant). These results were further verified from KEGG Mapper, Gorilla, and DAVID databases^{44–47}. The ontology from the STRING database indicated that 28 out of 38 proteins are involved in metabolic pathways and 10 of them were found in carbohydrates metabolism. The detailed ontology of the proteins has been enlisted in the Supplementary Table S4.

Next, the average clustering coefficient ($\langle CC \rangle$) of the network N was 0.5, signifying intermediate connectivity among the neighbors of the nodes and the presence of complete subgraphs or clique structures^{48,49}. There were three nodes with clustering coefficient (CC) as one i.e., all neighbors connected to each other, forming cliques. Cliques are essential as they are building blocks and the backbone of the network⁵⁰. Clique structures in the network make the network robust⁵¹ and stable⁵² as they are also

involved in the evolutionary process of the network from one stage to the other^{53,54}. The three proteins having $CC=1$ were COX6B1, MYH14, and SLC25A4 (Table 2). The proteins were further studied for their biological function through available literature, revealed their involvement in cardiac muscle contraction, neuropathy, myopathy, and oxidative phosphorylation. The protein COX6B1 is involved in cardiac muscle contraction⁵⁵. This protein encodes an integral, nuclear-encoded COX subunit⁵⁵. Mutation in the gene corresponding to COX6B1 protein, where histidine is replaced by cysteine at R20 residue, leads to reduced expression of COX6B1 protein in muscle and fibroblasts causing hypertrophic cardiomyopathy or cardiac muscle dysfunction⁵⁵. An autosomal dominant mutation in the protein MYH14 leads to neuropathy, myopathy, hoarseness, and hearing loss^{56,57}. SLC25A4 is active in cardiac hypertrophy and myopathy after mutation⁵⁸. It is also

Table 2. Potential network targets.

Serial no.	Protein	Metabolite	Similar drugs
1.	ACADM	Guanylic acid	DB03147
2.	COX4I1	Cholic acid, L-methionine	DB02659, DB04464
3.	COX6B1	Cholic acid, L-methionine	DB02659, DB04464
4.	HBB	Gallic acid, cholic acid	DB08262, DB07645
5.	MYH14	Guanylic acid	DB03126
6.	SLC25A4	Guanylic acid	DB00171

The proteins identified as potential targets from the network analysis are listed. The corresponding AR metabolites targeting these proteins with chemically and structurally similar drugs (DrugBank accession IDs) are listed.

reported that SLC25A4 controls oxidative phosphorylation, specifically to regulate energy phosphate levels associated with low ATP demand. This is directly associated with cardiac hypertrophy⁵⁸. These proteins are having a vital position in the network, vital in disease prognosis, can therefore be suggested as a probable drug targets.

Further, the betweenness centrality (β_c) of the nodes in the network signifies the participation of nodes in multiple pathways, indicating a positive correlation between the β_c of a node with its degree^{59,60}. A node with a low degree and high betweenness centrality is termed as nonhub bottlenecks (H_N). This parameter is crucial as it detects the nodes with occurrence in the maximum number of pathways. The targeting of high degree nodes may collapse the whole system as they are maximally connected. Therefore, such nonhub bottlenecks (H_N) are important for the network as they help to identify a weak breaking point of the network without compromising the whole system⁶¹. Three proteins ACADM, COX4I1, and HBB were identified in the nonhub bottleneck regime (Table 2), which are functionally involved in fatty acid metabolism, cardiac muscle contraction, and regulation of blood pressure. ACADM protein is associated with fatty acid metabolism⁶². Fatty acids and associated lipids are important determinants of both structure and function of cardiomyocytes. There is considerable evidence that in the postnatal and adult mammalian heart, fatty acid β oxidation is the preferred pathway for the energy that is required for efficient cardiac pumping^{62,63}. A major switch in the myocardial bio-energetic substrate used, from fatty acid to glucose leads to downregulation of fatty acid β enzyme, leading to cardiac failure⁵⁵. COX4I1 protein such as COX6B1 has a significant role in cardiac muscle contraction and its defect can result in heart failure, whereas HBB is responsible for the regulation of blood pressure⁵⁵. The COX4I1 protein, in case of myocardial insufficiency (heart failure) and dilated cardiomyopathy, is found to have decreased expression of COX4I1, which results in an impaired cytochrome c oxidase (CytOx) activity that has an important role in myocardial respiration, ultimately affecting the mitochondrial respiratory chain^{64,65}. Higher enzymatic activity but equal oxygen consumption contribute to the pathophysiology of myocardial insufficiency, which is an indicator of oxidative stress. The HBB protein has a vital role in the iron (Fe^{2+}) binding and oxygen-binding during the transport of blood in the circulatory system, which regulates blood pressure and blood vessel diameter in the body⁶⁶. Thus, these proteins can have therapeutic potential in treating cardiovascular diseases.

Combinatorial effects of the AR metabolites

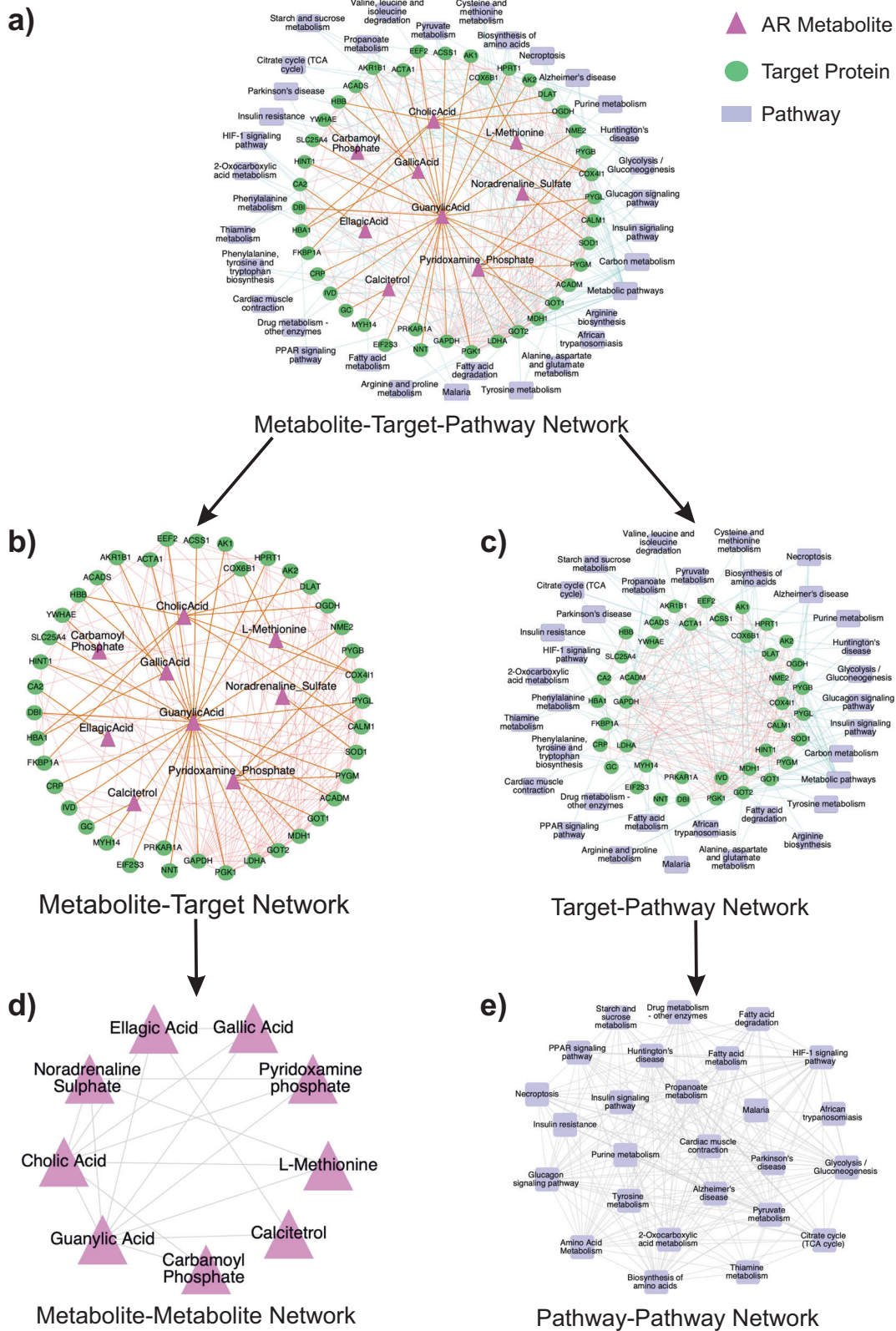
The proteins described in the PPI network discussed above were associated with nine of the 18 AR constituent metabolites (Supplementary Table S5). The interconnectedness of the PPI network indicates that these metabolites have a synergistic action,

which we analyzed through the guilt-by-association approach⁴. Multi-component or multilayered networks have been previously used for determining the synergistic drug action⁴. The basic principle involved in determining a synergistic action or synergistic partners is that two nodes of a type (metabolite/protein/pathway) should be connected through either a common or neighboring set of nodes of another kind. Therefore, to determine the synergy among AR metabolites, their target proteins, and the associated pathways; we constructed a multi-component network comprising AR metabolites, their potential protein targets, and associated pathways aka metabolite-target-pathway (MTP) network (Fig. 3a). This MTP network was used to construct two bilayer networks, Metabolite-Target (MT) network and Target-Pathway (TP) network (Fig. 3b, c).

The synergy at the metabolite level was determined by constructing a subnetwork from the constructed MTP network using metabolites and target proteins only (MT network, Fig. 3b). For any pair of metabolites to be synergistic partners, they should target the same or directly connected targets in order to affect a single or set of related pathways. Therefore, using a path length cutoff of ≤ 3 between two metabolites in the MT network for a direct connection, we created a metabolite-metabolite (MM) network, wherein directly connected nodes are synergistic couples (Fig. 3d). The analysis revealed that guanylic acid was directly connected to all the remaining metabolites (eight neighbors), indicating that guanylic acid is the most active synergistic ingredient of AR followed by cholic acid (seven neighbors). The other metabolites had 2–4 direct neighbors, and the MM network was a closed graph. This suggests that all the AR metabolites are directly or indirectly synergistic partners, and present themselves as a collective network of medicine.

To analyze the synergistic effect of the related pathways in the pharmacological action of AR metabolites, we used the target-pathway (TP) part (Fig. 3c) of the MTP network. A network medicine essentially affects multiple pathways simultaneously (co-regulation of pathways), in contrast to the traditional one molecule-one target-one pathway mode of drug action^{6,7}. Therefore, to study the co-regulation of different pathways by AR metabolites, we reduced the TP network to a pathway-pathway (PP) network, by using a path length cutoff of ≤ 3 in the TP network among pathway node sets to construct a direct connection between nodes in the PP network. The resulting PP network (Fig. 3e) is representative of different pathways co-regulated by the synergistic action of AR metabolites. PP network was observed to be a closed network comprising of 26 nodes (pathways) and 218 edges. The network being closed suggests high density which is evident from the average degree $\langle k \rangle$ of this network that comes out to be 17. Another topological feature namely diameter (D) for this network, was 2, indicating faster signaling⁴⁰. Further, a very high average clustering coefficient ($\langle CC \rangle$) of 0.9 signifies the presence of complete subgraphs, or in other words, it indicates that the pathways are interconnected. These network characteristics suggest that the co-regulation of all the discussed pathways in the PP network is critical to disease treatment and therefore, the removal of even one of the nine metabolites from the AR concoction could result in the disruption of pharmacological action, thereby, reaffirming that the AR metabolites have a synergistic response.

To understand the synergy of drug action at the target level, we investigated the protein-protein interaction network. The synergistic effect of metabolites in this PPI network was studied through the interconnectedness among the target proteins of different metabolites. The PPI network created is part of a larger disease-specific PPI network, which is targeted by the metabolites. Therefore, the network (Fig. 4) is vital for disease treatment and the collapse of this network would lead to a disruption of drug action. The knockout or deletion of individual metabolite-associated target proteins in the network will disrupt the whole



network. Thus, we categorized the proteins in accordance with their corresponding metabolites (Fig. 4). We observed that various proteins associated with different metabolites are vital for maintaining network integration. This was evident from the proteins associated with guanylic acid, as they are maximally

present in the underlined network, further suggesting that this metabolite has significant involvement in response to disease treatment. The other metabolites such as cholic acid, gallic acid, ellagic acid, etc. associated proteins provide a bridge of connections for the connectivity of guanylic acid-associated

Fig. 3 Synergistic action of AR metabolites. **a** Metabolite-Target-Pathway (MTP) network was constructed by connecting metabolites with their corresponding target proteins in the protein-protein interaction (PPI) network. The target proteins were further connected with their associated pathways to produce the final MTP network. The metabolites are shown as pink triangles, target proteins are represented as green circles, and pathways as blue rectangles. **b** The metabolite-target protein interaction network (MT network) is the subgraph of the MTP network and consists of the interaction of metabolites with their corresponding target proteins and interactions among the target proteins. **c** The target protein-pathway interaction network (TP network) consists of the protein-protein interactions and protein-pathway relations. The proteins are shown in green, while the pathways are shown in blue. **d** MT network was further reduced to a metabolite-metabolite interaction network (MM network) using a distance cutoff of ≤ 3 in the MT network for each metabolite pairs to identify a direct connection in the MM network. The direct association between two metabolites signifies their co-action in the regulation of single or multiple pathways. **e** TP network was reduced to a pathway-pathway interaction network (PP network). The PP network was constructed using the same principle used in the construction of the MM network. The adjacency (direct connection) of two nodes signifies the co-regulation of two pathways due to the action of Amalaki Rasayana metabolites.

proteins among each other and in other pathways. This observation signifies the utility of multiple bridging metabolites for the therapeutic action of guanylic acid. This behavior was analyzed by studying the ontology of these proteins that revealed the interrelation of the proteins in the regulation of multiple pathways, such as metabolic pathways, glycolysis/gluconeogenesis, insulin signaling pathway, fatty acid metabolism, cardiac muscle contraction, etc. (Supplementary Table S4).

The three levels of synergistic action investigated, thus, reveal that the action of AR constituents is combinatorial and reflective of the modern concept of network medicine.

Drug repositioning for pressure-overload LVCH

The six proteins ACADM, COX4I1, HBB, COX6B1, MYH14, and SLC25A4 correspond to four metabolites namely guanylic acid, cholic acid, L-methionine, and gallic acid. These metabolites may have therapeutic effects in the treatment of pressure-overload LVCH. Therefore, we studied the biological functions of these candidate metabolites for therapeutic implications and verified their binding affinity with the target proteins. Guanylic acid is an experimental drug and mechanisms in which this metabolite involved are unknown. This metabolite belongs to a class of organic compound ribonucleoside 3'-phosphates that contain a phosphate group attached to the C-3 carbon of the ribose or deoxyribose unit. It is known to be involved in uracil phosphoribosyltransferase activity⁶⁷. Cholic acid is an approved drug that has a major role in primary bile acid production in the liver. It facilitates fat absorption and cholesterol excretion. This metabolite is primarily used in the treatment of children and adults with bile acid synthesis disorders and for peroxisomal disorders (such as Zellweger syndrome—bile enzyme malfunction)^{68,69}. L-methionine is an approved drug and nutraceutical that helps to lower cholesterol levels by increasing the production of lecithin in the liver³⁴. It reduces liver fat, protects the kidneys, and prevents disorders relating to hair, skin, and nails. The metabolite is a sulfur-based essential amino acid that acts as a natural chelating agent for heavy metals and also regulates the production of ammonia in urine⁷⁰. Gallic acid is an approved drug that possesses protective effects on the gastric mucosa and has strong astringent effects. It is used as a deodorizing agent in medicines. The metabolite plays a major role in forming a protective coat on the intestinal mucosa and treating ulcers e.g., ulcers from *H. pylori*⁷¹. It has antimicrobial effects against various gastrointestinal tract pathogens. It has reported similarities with ellagic acid, which helps in blood clotting and subsequently reduce bleeding⁷¹. Further, all these metabolites are reported to be directly absorbed in plasma while administered orally (Supplementary Table S6). After examining the functional significance of the metabolites corresponding to proteins obtained from network analysis, we characterized their biophores to propose repositioning of these metabolites against available drugs.

To determine the set of existing drugs for repositioning in the treatment of pressure-overload LVCH, we performed molecular docking of a set of proteins obtained from network analysis for drug-target protein and metabolite-target protein pairs. The similarity of the biophores for the two pairs of complexes would suggest the interchangeability among available drugs and AR metabolites for therapeutic applications^{72,73}. We observed similar biophores for the selected set of 6 target proteins viz. ACADM, COX4I1, HBB, COX6B1, MYH14, and SLC25A4 (Table 2, Supplementary Figs. S2–S7, and Supplementary Table S7). The COX4I1 and COX6B1 corresponded to identical drug and metabolite complexes and individually were found to form similar biophores. However, the protein HBB had two metabolites and corresponding similar drugs. Out of the two, one of the biophores obtained for HBB, namely HBB-cholic acid and HBB-sebacic acid (DrugBank Id: DB07645) complexes were unidentical and hence ignored for further study (Supplementary Fig. S8).

The similar biophores suggest similar biological effects of AR metabolites and their structurally similar drugs (Supplementary Section I and Supplementary Table S8), which will further be investigated for their re-purposing to treat pressure-overload LV cardiac hypertrophy in future studies. Further, we briefly reviewed the biological activity of protein-drug complexes having similar biophores to the protein-AR metabolite complexes, because the information could be useful for the use of the AR metabolites to be potential alternatives to the currently used drugs. The detailed information on the mechanism of these pathways is summarized here.

The first protein-drug complex ACADM-flavin adenine dinucleotide (DrugBank Id: DB03147) complex has a significant role in the Lovastatin pathway and Cerivastatin pathway that are involved in the lowering and inhibition of cholesterol synthesis respectively⁷⁴. The proteins COX4I1 and COX6B1 are part of identical drug and metabolite complexes namely (COX4I1-cholic acid (DrugBank Id: DB02659))—(COX4I1-N-formyl methionine (DrugBank Id: DB04464)), and (COX6B1-cholic acid)—(COX6B1-N-formyl methionine), respectively. The complexes play an important role in cytochrome c oxidase activity and regulate bile acid synthesis as well as in the initiation of protein synthesis⁷⁵. The protein-drug complex HBB-dicarboxy naphthalene (DrugBank Id: DB08262) regulates oxygen transporter activity thereby decreasing blood pressure in the body⁷⁶. Next, MYH14-Mant-ADP (DrugBank Id: DB03126) complex helps in Microfilament motor activity that has a major role in myopathy and cytokinesis⁷⁷. The last complex SLC25A4-Adenosine Tri Phosphate (DrugBank Id: DB00171) has a crucial role in metabolism and is also a neurotransmitter. It is involved in adenine transmembrane transporter activity which catalyzes the exchange of cytoplasmic ADP with mitochondrial ATP, across the mitochondrial inner membrane⁷⁸. All these complexes were found to be inhibitors, indicating their role in suppressing one or the other pathways.

Considering the fundamental principles of similar structure similar function^{31,32}, we propose that the metabolites namely

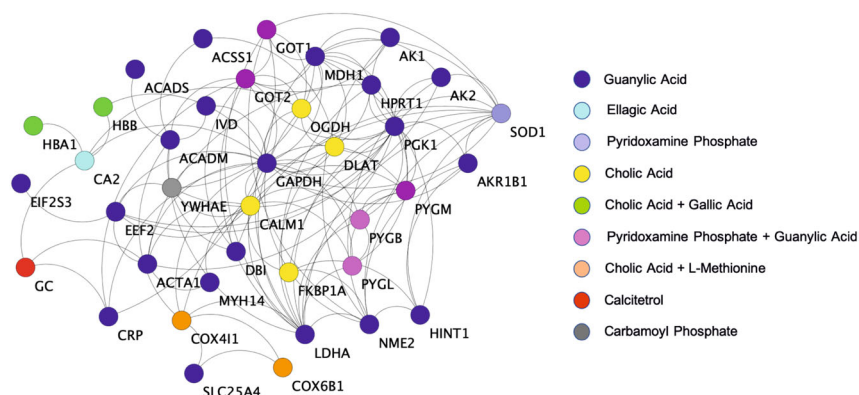


Fig. 4 Synergistic action of AR metabolites on target proteins. The color codes on the proteins/targets in the interaction network indicate their corresponding metabolites identified after screening using SMSD and Tc. Blue circles represent the proteins corresponding to guanylic acid (the maximum in number), followed by yellow color for cholic acid. ellagic acid is represented by cyan color, whereas calcitriol by red circles. Gray color circles correspond to carbamoyl phosphate and purple to pyridoxamine phosphate. Noradrenaline sulfate metabolite corresponds to light blue circled proteins. Few proteins correspond to more than one metabolite and are represented as orange circled proteins for the combination of cholic acid and L-methionine, and green circles for cholic acid and gallic acid. The pink circles are proteins corresponding to pyridoxamine phosphate and guanylic acid. All these proteins are interconnected and have a significant role in the disruption or regulation of pathways as seen when knocked out from the network. This observation suggests the co-dependency of AR metabolites for the maintenance of the protein–protein interaction network, which is vital to drug action, thereby affirming their synergistic action.

guanylic acid, cholic acid, L-methionine, and gallic acid corresponding to the protein targets ACADM, COX411, COX6B1, HBB, MYH14, and SLC25A4 can be better alternatives to the known drugs as they are extracted from a natural source and have minimal side effects.

DISCUSSION

The findings of our study indicate the scope for revitalizing Ayurvedic concoctions to be used as therapeutic agents. Using the integrated approach of cheminformatics and network systems biology, we investigated AR, an Ayurvedic rejuvenate for the treatment of cardiovascular diseases (pressure-overload LVCH). The authentic text for the preparation and functional significance of AR is Charaka Chikitsa Sthana, written in the third century BCE. Most of the multi-component formulations in Ayurveda were believed to be acting on multiple targets and therefore, demands a “systems approach”. Such a systematic study could not have been possible without contributory support from in-vivo studies, proteomics analysis, informatics tools, and techniques of network pharmacology. Through this work, we made a conscious attempt to present a complete analytical platform to investigate the efficacy and synergy of ayurvedic medicines systematically.

The in-vivo study on artificially Aorta Constricted and aged Wistar rats provided vital information about changes in protein expression profile, yielding 450 (List L1) and 1166 (List L2) proteins, respectively, that are modulated by AR. The chemical and structural similarity analysis identified structurally similar drugs to AR metabolites, with a turnout of 1004 proteins (List L3). The intersection of protein expression data and structural analysis data (List L4) was used together to construct the PPI network using annotations from existing databases. The network of common proteins was further analyzed for its topological properties such as degree, clustering coefficient, betweenness centrality, etc. The clustering coefficient reveals COX6B1, MYH14, and SLC25A4 proteins that form clique structures or complete subgraphs in the network. The cliques are essential as they form the backbone of a network⁵⁰ and are preserved structures that are responsible for the evolution of any system. They are known as responsible for the robustness⁵¹ and stability⁵² of the underlying system. Further, the proteins corresponding to nonhub bottlenecks (H_N) i.e., nodes having a low degree and high betweenness centrality regime are

ACADM, COX411, and HBB. The proteins with H_N identify the weak breaking points in the network as they are involved in many pathways. The biological functions of these proteins signify either a regulatory or pathogenic role in the cardiovascular system, such as blood pressure regulation, fatty acid synthesis, cholesterol synthesis, cardiac hypertrophy, cardiomyopathy, and systemic hypertension.

We also constructed and analyzed the metabolite–target–pathway network to determine the synergistic action of AR. The synergy was evaluated qualitatively using the guilt-by-association method⁴ at the metabolite, pathway, and target levels individually to ascertain the combinatorial effect of AR metabolites. Next, the six proteins from network analysis were back-traced to their corresponding AR metabolites and similar drugs. We identified that these existing drug molecules can be repositioned to develop a combinatorial therapy for the treatment of pressure-overload LV cardiac hypertrophy and other cardiovascular diseases.

Though this study details the synergistic action of AR constituents and provides information on the constitution of a repositioned multidrug formulation for the treatment of pressure-overload LVCH, it is limited by the quantitative composition of the repositioned drugs. A possible limitation of this study is the bioavailability of AR metabolites in plasma upon oral administration. However, the plasma absorbance of cholic acid, gallic acid, and L-methionine has been previously reported (Supplementary Table S6). The clinical applications including drug–dose relationship, efficacy considerations, and toxic effects of such drugs need to be rigorously confirmed. A quantitative ratio of the repositioned drugs or even AR constituent metabolites would be required for the further development of combinatorial network medicine. This is the point of exploration for our future studies for developing a therapy for the treatment of pressure-overload LV cardiac hypertrophy, using existing drug molecules and AR constituent metabolites. The analytical approach presented here can be extended further to various other diseases as well. This framework of employing tools of cheminformatics and network biology may provide a direction for developing new drugs, therapeutic targets, biomarkers, and repositioning existing drugs for complex diseases in a time-efficient and cost-effective manner.

METHODS

LC-MS of AR concoction resulted in the identification of 18 metabolites, as significant constituents of AR. The in-vivo experiments provided information containing details of proteins extracted from the proteome analysis (LC-MS) of rat models. The proteome data were recorded for the test (AR treated), placebo or carrier (Ghee-Honey treated), and the control (Untreated) samples in Aorta Constricted and Biologically Aged rats. The 18 metabolites were screened for similar drugs by comparing the structure of AR metabolites with that of drug molecules available in the DrugBank database, employing SMSD and Tc as screening methods. The drugs which are similar to AR metabolites and their associated drug targets were named in List L3. The common proteins among targets from in-vivo study and similarity search (intersection of L1, L2, and L3) resulted in List L4. The PPI network analyses for these common proteins were examined for their network and pathway integrity. Further, we determined the synergy among AR metabolites, their target proteins, and the associated pathways by constructing a multi-component network employing the guilt-by-association method⁴. This network comprised of AR metabolites, their potential protein targets, and associated pathways (MTP network)⁴. The proteins obtained from network analysis were also scrutinized for re-purposing by characterizing their respective biophores, using molecular docking. Detailed procedures employed in this study are the following.

in-vivo study: data source and acquisition

The in-vivo study of AR on pressure-overload LV cardiac hypertrophy²⁸ comprised of the proteomic information of two sets of male Wistar rats (i) AC with pressure-overload LVCH, induced by clipping ascending aorta with titanium clips, and (ii) BA for age-associated cardiac dysfunction. The technique for proteome expression analysis LC-MS was based on the principle that a protein is only detected when it is present above a threshold. The absence of proteins in the control conditions represents the downregulation of those proteins in the disease model. Conversely, the presence of the same proteins in the test condition indicates the upregulation of the proteins due to AR treatment. The objective of our study was to identify proteins, which are directly targeted by the metabolites present in Amalaki Rasayana (AR). Therefore, for our study, we considered all the proteins that are detected from the LC-MS analysis for both the AC and BA rats, resulting in List L1 and L2, respectively. The proteomics analysis had raw files containing details of proteins expressed in all the three sample replicates (AR Treated, Placebo, and Untreated) as RefSeq accessions along with their description, expression scores, etc. For our analysis, we considered only the RefSeq IDs and the expression scores for each sample replicates. Thereafter, we averaged over their expression scores as a representative value for all the replicates (Supplementary Dataset S1). Further, we extracted and converted these RefSeq accessions to protein UNIPROT IDs to deduce the number of proteins for each BA and AC groups as L1 and L2 protein lists (Supplementary Dataset S1).

The in-vivo study was performed at Rajiv Gandhi Center for Biotechnology with the approval of the Institutional animal ethics committee (IAEC) in Rajiv Gandhi Center for Biotechnology (RGCB) under protocol no. IAEC/150/CCK/2012, strictly following the rules and regulations of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA), Government of India. The results of this experiment, however, have already been published²⁸. In this manuscript, we have only used the proteomics data of the earlier study, and therefore in-vivo studies are not in principle, part of this paper.

Drug-metabolite similarity search and ontology

The protein targets associated with these drugs (similar to AR metabolites) were considered the representative set of proteins that the AR metabolites can interact with. This is based on the universal philosophy of drug discovery, like structure leads to like functions. To identify potential drug targets for the AR metabolites, we compared the structure of AR metabolites with the structures of well-known modern drug molecules available in the DrugBank database³⁴. The similarity index, namely the Tc between drug molecules and metabolites (from AR) were generated using Open Babel⁷⁹. Tc is defined as the ratio of the intersecting set of a molecule to the union set calculated as the measure of similarity³⁵. Mathematically, Tc can be represented as $Tc(a, b) = N_c / (N_a + N_b - N_c)$, where N is the number of attributes in each molecule (a, b) and C is the intersection set.

The pairwise similarity visualization of drug molecules and the AR metabolites was done using SMSD toolkit³⁶. Protein ontologies were obtained from KEGG Mapper, Gorilla, and DAVID databases^{44–47}.

Construction and analysis of protein–protein interaction networks

This representative set of proteins play a significant role in therapeutics as they were altered by the treatment with AR as well as targeted corresponding to AR metabolites. Networks can be represented as a collection of nodes connected to their edges. Protein–protein interactions (PPI) from the STRING database were used for the construction of interaction map/network⁸⁰. The P value of the interactions was retrieved from the STRING database, which uses a hypergeometric test for statistical significance of the data⁸⁰. Further, the interactions were reconstituted into binary adjacency matrix (A) where two proteins, if interacting, were denoted as 1 and 0 if otherwise. Thereafter, topological features, which include structural properties of the network such as degree, clustering coefficient, betweenness centrality, diameter, etc. were calculated. These may reveal the minimal set of drug targets for cardiovascular diseases.

The fundamental topological parameter in a network is the degree of a node (k_i), which is defined as the number of first neighbors the node has ($k_i = \sum A_{ij}$)⁵⁹. Another critical parameter is the clustering coefficient (CC) of the network. The clustering coefficient of a node i (CC_i) is defined as the ratio of the number of connections a particular node has and the possible number of connections the particular node can have⁴⁹, as described by Eq. (1).

$$CC_i = \frac{2 \sum_{j_2=1}^{k_i} \sum_{j_1=1}^{k_i} (A_{ij_1} A_{ij_2} A_{j_1 j_2})}{k_i(k_i - 1)} \quad (1)$$

where i is the node of interest and j_1 and j_2 are any two neighbors of the node i and k_i are the degrees of the node i . The average clustering coefficient ($\langle CC \rangle$) of a network is described by Eq. (2)

$$CC = \frac{1}{n} \sum_{i=1}^n CC_i \quad (2)$$

These are also known as cliques. They are complete subgraphs in the network which are known to be the conserved part of the network⁵⁰. The average clustering coefficient of the network characterizes the overall tendency of nodes to form clusters or groups. Further, the diameter of the network measures the longest of the shortest path between all the pairs of nodes⁴⁰. It gives information on how fast a signal can transmit to the whole network.

Next, the betweenness centrality (β_c) of a node i is defined as the fraction of the shortest paths between node pairs that pass through the said node of interest⁵⁹, as described by Eq. (3)

$$\beta_{ci} = \sum_{st} \frac{n_{st}^i}{g_{st}} \quad (3)$$

where, n_{st}^i is the number of paths from s to t that passes through i and g_{st} is the total number of paths from s to t in the network.

Synergistic effects of AR constituents

The use of drug combinations against multiple linked targets in a disease-specific network is a more sensible approach for the treatment of complex diseases. Ayurveda is the science of combinatorial therapy and its medicines often have multiple constituents, which are considered to operate through a cooperative mechanism termed synergy. Most of the research on combinatorial medicines has been based on the hypothesis that the synergistic action of drugs is due to their combinatorial effects on targets identified by their network topological features. Thus, we categorized the proteins in accordance with their corresponding metabolites and observed the presence of proteins and their corresponding metabolites and pathways in the maintenance of network integration. We use the guilt-by-association approach⁴ where the distance of ≤ 3 among the connected entities was taken into consideration for each dataset.

Identification of binding sites: molecular docking for re-purposing

The proteins obtained from network analysis were back-traced to their metabolites and drug targets for identification of the binding sites among the protein–metabolite and protein–drugs through molecular docking

using AutoDock Vina⁸¹. To perform this exercise, the structures of the proteins were identified from Protein Data Bank (PDB)⁸². Thereafter, the selected proteins retrieved from PDB were using Chimera separated from the ligands and chains that are not relevant to the present study. The water molecules were also removed by using AutoDockTools (ADT)⁸¹. This was followed by the addition of Gasteiger charges and adjusting the charges on each residue to bring integral charges across the macromolecule. Hydrogen atoms were added, the non-polar ones were merged, and the macromolecule thus obtained was saved as a PDBQT file. Next, the ligand structures were retrieved from DrugBank and PubChem databases³⁴. They were energy minimized using Avogadro's 1.2.0 steepest descent algorithm to reduce any bad conformations that may be introduced and give artefacts of potential energy. It is likely that these conformations were not preserved when docking was finally carried out since the ligand is treated as a flexible entity in AutoDock Vina. It was read into ADT, torsion tree root detected, and processed as another PDBQT file. This modified structure (preparatory files) was then further docked blindly.

The strategy behind using blind docks relied on the fact that we can scale the exhaustiveness with grid box and allow the docking to be completed faster, along with generally better results utilizing the system's multithreaded architecture⁸¹. The docking was run and examined using LigPlot and PyMol for screening the biophores of the docked protein and ligand (drug/metabolite)⁸³. Further, we generated a biophore fingerprint table to identify similar (recurring) residues in the binding pocket of protein–metabolite and protein–drug complexes.

Reporting summary

Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

DATA AVAILABILITY

All the relevant data supporting the findings of this study are included in the paper and its Supplementary Information files. Processed data in the form of figures and tables are presented in the main paper. The Supplementary Information contains two files: (1) Supplementary Information file with biophore details, network topologies, and information related to identified targets. (2) Supplementary Dataset S1: This file has details of target profiling of AC and BA rat models, Tanimoto coefficient (Tc) of drugs with respect to their similar metabolites, common protein expression data, etc.

CODE AVAILABILITY

For this study, no custom made codes and algorithms were used.

Received: 3 August 2020; Accepted: 13 January 2021;

Published online: 15 February 2021

REFERENCES

- Rawlins, S. M. The rise & fall of modern medicine. *Nat. Med.* **5**, 865–865 (1999).
- Kumar, A., Tiwari, A. & Sharma, A. Changing paradigm from one target one ligand towards multi-target directed ligand design for key drug targets of Alzheimer disease: an important role of in silico methods in multi-target directed ligands design. *Curr. Neuropharmacol.* **16**, 726–739 (2018).
- Reddy, A. S. & Zhang, S. Polypharmacology: drug discovery for the future. *Expert Rev. Clin. Pharm.* **6**, 41–47 (2013).
- Casas, A. I. et al. From single drug targets to synergistic network pharmacology in ischemic stroke. *Proc. Natl Acad. Sci. USA* **116**, 7129–7136 (2019).
- Castillo-Quan, J. I. et al. A triple drug combination targeting components of the nutrient-sensing network maximizes longevity. *Proc. Natl Acad. Sci. USA* **116**, 20817–20819 (2019).
- Barabási, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68 (2011).
- Sonawane, A. R., Weiss, S. T., Glass, K. & Sharma, A. Network medicine in the age of biomedical big data. *Front. Genet.* **10** <https://doi.org/10.3389/fgene.2019.00294> (2019).
- Cheng, F. et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **9**, 2691 (2018).
- Zhang, R., Zhu, X., Bai, H. & Ning, K. Network pharmacology databases for traditional Chinese medicine: review and assessment. *Front. Pharmacol.* **10** <https://doi.org/10.3389/fphar.2019.00123> (2019).
- Wu, Z. et al. In silico prediction of chemical mechanism of action via an improved network-based inference method. *Br. J. Pharmacol.* **173**, 3372–3385 (2016).
- Gu, J. et al. Use of natural products as chemical library for drug discovery and network pharmacology. *PLoS ONE* **8**, e62839 (2013).
- Spirito, P. et al. Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy. *N. Engl. J. Med.* **342**, 1778–1785 (2000).
- Weissler-Snir, A. et al. Hypertrophic cardiomyopathy-related sudden cardiac death in young people in Ontario. *Circulation* **140**, 1706–1716 (2019).
- Jadczyk, T., Faulkner, A. & Madeddu, P. Stem cell therapy for cardiovascular disease: the demise of alchemy and rise of pharmacology. *Br. J. Pharmacol.* **169**, 247–268 (2013).
- Wong, K., Boheler, K. R., Petrou, M. & Yacoub, M. H. Pharmacological modulation of pressure-overload cardiac hypertrophy. *Circulation* **96**, 2239–2246 (1997).
- Toepfer, C. N. et al. Hypertrophic cardiomyopathy mutations in MYBPC3 dysregulate myosin. *Sci. Transl. Med.* **11**, eaat1199 (2019).
- Frey, N., Katus, H. A., Olson, E. N. & Hill, J. A. Hypertrophy of the heart. *Circulation* **109**, 1580–1589 (2004).
- Samak, M. et al. Cardiac hypertrophy: an introduction to molecular and cellular basis. *Med. Sci. Monit. Basic Res.* **22**, 75–79 (2016).
- Marian, A. J. & Braunwald, E. Hypertrophic cardiomyopathy: genetics, pathogenesis, clinical manifestations, diagnosis, and therapy. *Circulation Res.* **121**, 749–770 (2017).
- Chauhan, A., Semwal, D. K., Mishra, S. P. & Semwal, R. B. Ayurvedic research and methodology: present status and future strategies. *Ayu* **36**, 364–369 (2015).
- Mukherjee, P. K., Banerjee, S. & Kar, A. Exploring synergy in ayurveda and traditional Indian systems of medicine. *Synergy* **7**, 30–33 (2018).
- Parasuraman, S., Thing, G. S. & Dhanaraj, S. A. Polyherbal formulation: concept of ayurveda. *Pharmacogn. Rev.* **8**, 73–80 (2014).
- Zhang, C. & Yan, G. Synergistic drug combinations prediction by integrating pharmacological data. *Synth. Syst. Biotechnol.* **4**, 67–72 (2019).
- Celebi, R., Bear Don't Walk, O., Movva, R., Alpsy, S. & Dumontier, M. In-silico prediction of synergistic anti-cancer drug combinations using multi-omics data. *Sci. Rep.* **9**, 8949 (2019).
- Chen, D., Zhang, H., Lu, P., Liu, X. & Cao, H. Synergy evaluation by a pathway–pathway interaction network: a new way to predict drug combination. *Mol. Biosyst.* **12**, 614–623 (2016).
- Shen, C.-Y., Jiang, J.-G., Yang, L., Wang, D.-W. & Zhu, W. Anti-ageing active ingredients from herbs and nutraceuticals used in traditional Chinese medicine: pharmacological mechanisms and implications for drug discovery. *Br. J. Pharmacol.* **174**, 1395–1425 (2017).
- Hopkins, A. L. Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* **4**, 682–690 (2008).
- Kumar, V. et al. Amalaki rasayana, a traditional Indian drug enhances cardiac mitochondrial and contractile functions and improves cardiac function in rats with hypertrophy. *Sci. Rep.* **7**, 8588 (2017).
- Acharya, J. T. Charaka Samhita of Agnivesha. Edition 1. (Chaukhamba Sanskrit Sansthan. Varanasi, 2011).
- Forsberg, E. M. et al. Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS online. *Nat. Protoc.* **13**, 633–651 (2018).
- Mohd Fauzi, F. et al. Chemogenomics approaches to rationalizing the mode-of-action of traditional chinese and ayurvedic medicines. *J. Chem. Inf. Modeling* **53**, 661–673 (2013).
- Fauzi, F. M. et al. Linking ayurveda and western medicine by integrative analysis. *J. Ayurveda Integr. Med* **4**, 117–119 (2013).
- Rognan, D. Chemogenomics approaches to rational drug design. *Br. J. Pharmacol.* **152**, 38–52 (2007).
- Wishart, D. S. et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **36**, D901–D906 (2008).
- Bajusz, D., Rácz, A. & Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* **7**, 20–20 (2015).
- Rahman, S. A., Bashton, M., Holliday, G. L., Schrader, R. & Thornton, J. M. Small molecule subgraph detector (SMSD) toolkit. *J. Cheminform.* **1**, 12–12 (2009).
- Goh, K.-I. et al. The human disease network. *Proc. Natl Acad. Sci. USA* **104**, 8685–8690 (2007).
- Rai, A. et al. Understanding cancer complexome using networks, spectral graph theory and multilayer framework. *Sci. Rep.* **7**, 41676 (2017).
- Rai, A., Menon, A. V. & Jalan, S. Randomness and preserved patterns in cancer network. *Sci. Rep.* **4**, 6368 (2014).
- Rai, A., Pawar, A. K. & Jalan, S. Prognostic interaction patterns in diabetes mellitus II: a random-matrix-theory relation. *Phys. Rev. E* **92**, 022806 (2015).
- Zhou, X., Menche, J., Barabási, A.-L. & Sharma, A. Human symptoms–disease network. *Nat. Commun.* **5**, 4212 (2014).
- Wu, J., Tan, S.-Y., Liu, Z., Tan, Y.-J. & Lu, X. Enhancing structural robustness of scale-free networks by information disturbance. *Sci. Rep.* **7**, 7559–7559 (2017).

43. min, B. & Zheng, M. Correlated network of networks enhances robustness against catastrophic failures. *PLoS ONE* **13**, e0195539 (2018).
44. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinforma.* **10**, 48 (2009).
45. Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
46. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
47. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2015).
48. Takemoto, K. & Osawa, C. Evolving networks by merging cliques. *Phys. Rev. E* **72**, 046116 (2005).
49. Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998).
50. Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits* (CRC Press, 2019).
51. Yeager-Lotem, E. et al. Network motifs in integrated cellular networks of transcription–regulation and protein–protein interaction. *Proc. Natl Acad. Sci. USA* **101**, 5934–5939 (2004).
52. Dwivedi, S. K. & Jalan, S. Emergence of clustering: role of inhibition. *Phys. Rev. E* **90**, 032803 (2014).
53. Adamcsek, B., Palla, G., Farkas, I. J., Derényi, I. & Vicsek, T. CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics* **22**, 1021–1023 (2006).
54. Li, J., Wang, X. & Cui, Y. Uncovering the overlapping community structure of complex networks by maximal cliques. *Phys. A: Stat. Mech. its Appl.* **415**, 398–406 (2014).
55. Abdulhag, U. N. et al. Mitochondrial complex IV deficiency, caused by mutated COX6B1, is associated with encephalomyopathy, hydrocephalus and cardiomyopathy. *Eur. J. Hum. Genet.* **23**, 159–164 (2015).
56. Hodatsu, A. et al. Impact of cardiac myosin light chain kinase gene mutation on development of dilated cardiomyopathy. *ESC Heart Fail.* **6**, 406–415 (2019).
57. Rinaldi, F. et al. Aberrant splicing and expression of the non muscle myosin heavy-chain gene MYH14 in DM1 muscle tissues. *Neurobiol. Dis.* **45**, 264–271 (2012).
58. Tossierams, A. et al. Two new cases of mitochondrial myopathy with exercise intolerance, hyperlactatemia and cardiomyopathy, caused by recessive SLC25A4 mutations. *Mitochondrion* **39**, 26–29 (2018).
59. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
60. Barabási, A.-L. & Oltvai, Z. N. Network biology: understanding the cell’s functional organization. *Nat. Rev. Genet.* **5**, 101–113 (2004).
61. Furlong, L. I. Human diseases through the lens of network biology. *Trends Genet.: TIG* **29**, 150–159 (2013).
62. Mozaffarian, D., Katan, M. B., Ascherio, A., Stampfer, M. J. & Willett, W. C. Trans fatty acids and cardiovascular disease. *N. Engl. J. Med.* **354**, 1601–1613 (2006).
63. Marín-García, J. & Goldenthal, M. J. Fatty acid metabolism in cardiac failure: biochemical, genetic and cellular analysis. *Cardiovascular Res.* **54**, 516–527 (2002).
64. Abu-Libdeh, B. et al. Mutation in the COX4I1 gene is associated with short stature, poor weight gain and increased chromosomal breaks, simulating Fanconi anemia. *Eur. J. Hum. Genet.* **25**, 1142–1146 (2017).
65. Vogt, S. et al. Myocardial insufficiency is related to reduced subunit 4 content of cytochrome c oxidase. *J. Cardiothorac. Surg.* **13**, 95 (2018).
66. Gouagna, L. C. et al. Genetic variation in human HBB is associated with Plasmodium falciparum transmission. *Nat. Genet.* **42**, 328–331 (2010).
67. Ghim, S. Y. & Neuhard, J. The pyrimidine biosynthesis operon of the thermophile *Bacillus caldolyticus* includes genes for uracil phosphoribosyltransferase and uracil permease. *J. Bacteriol.* **176**, 3698–3707 (1994).
68. Craddock, A. L. et al. Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter. *Am. J. Physiol.* **274**, G157–G169 (1998).
69. Fickert, P. et al. Effects of ursodeoxycholic and cholic acid feeding on hepatocellular transporter expression in mouse liver. *Gastroenterology* **121**, 170–183 (2001).
70. Tandon, S. K., Singh, S. & Flora, S. J. Influence of methionine and zinc supplementation during chelation of lead in rats. *J. Trace Elem. Electrolytes Health Dis.* **8**, 75–77 (1994).
71. Sharifi-Rad, M. et al. Anticancer agents: from plant extracts to phytochemicals in healing promotion. *Molecules* **23**, 1751 (2018).
72. Berger, S. I. & Iyengar, R. Role of systems pharmacology in understanding drug adverse events. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **3**, 129–135 (2011).
73. Ekins, S., Mestres, J. & Testa, B. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* **152**, 9–20 (2007).
74. Manzoni, M. & Rollini, M. Biosynthesis and biotechnological production of statins by filamentous fungi and application of these cholesterol-lowering drugs. *Appl. Microbiol. Biotechnol.* **58**, 555–564 (2002).
75. Krisans, S. K., Thompson, S. L., Pena, L. A., Kok, E. & Javitt, N. B. Bile acid synthesis in rat liver peroxisomes: metabolism of 26-hydroxycholesterol to 3 beta-hydroxy-5-cholenic acid. *J. Lipid Res.* **26**, 1324–1332 (1985).
76. Ali, A., Baby, B., Soman, S. S. & Vijayan, R. Molecular insights into the interaction of hemorphin and its targets. *Sci. Rep.* **9**, 14747 (2019).
77. Heissler, S. M. & Manstein, D. J. Comparative kinetic and functional characterization of the motor domains of human nonmuscle myosin-2C isoforms. *J. Biol. Chem.* **286**, 21191–21202 (2011).
78. von Renesse, A. et al. Muscle weakness, cardiomyopathy, and L-2-hydroxyglutaric aciduria associated with a novel recessive SLC25A4 mutation. *JIMD Rep.* **43**, 27–35 (2019).
79. O’Boyle, N. M. et al. Open Babel: an open chemical toolbox. *J. Cheminform.* **3**, 33 (2011).
80. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
81. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
82. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
83. Mehra, A., Jerath, G., Ramakrishnan, V. & Trivedi, V. Characterization of ICAM-1 biophore to design cytoadherence blocking peptides. *J. Mol. Graph. Model.* **57**, 27–35 (2015).

ACKNOWLEDGEMENTS

This work was supported by Department of Biotechnology project number BT/PR2526/NER/95/1238/2017 and BT/565/NE/U-EXCEL/2016. A.R. was financially supported by the Indian Institute of Technology Guwahati (Project ID: BSBE/IPDF/2018-19/APAR/01). A.R. thanks Amay Redkar for helping in the docking experiment exercise.

AUTHOR CONTRIBUTIONS

C.C.K. and V.R. designed the research and managed the project; A.R. performed research; V.K. contributed in conducting wet-lab experiments and generating primary research data; G.J. analyzed the data; and A.R. wrote the paper. All the authors contributed to the improvement of the paper.

COMPETING INTERESTS

Filing of two separate patent applications has been initiated for IP protection by IIT Guwahati (TEMP/E1/28937/2020-KOL and TEMP/E1/28939/2020-KOL, Patent office Kolkata, India).

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41540-021-00171-z>.

Correspondence and requests for materials should be addressed to C.C.K. or V.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021