Next-Generation Sequencing Data Analyses

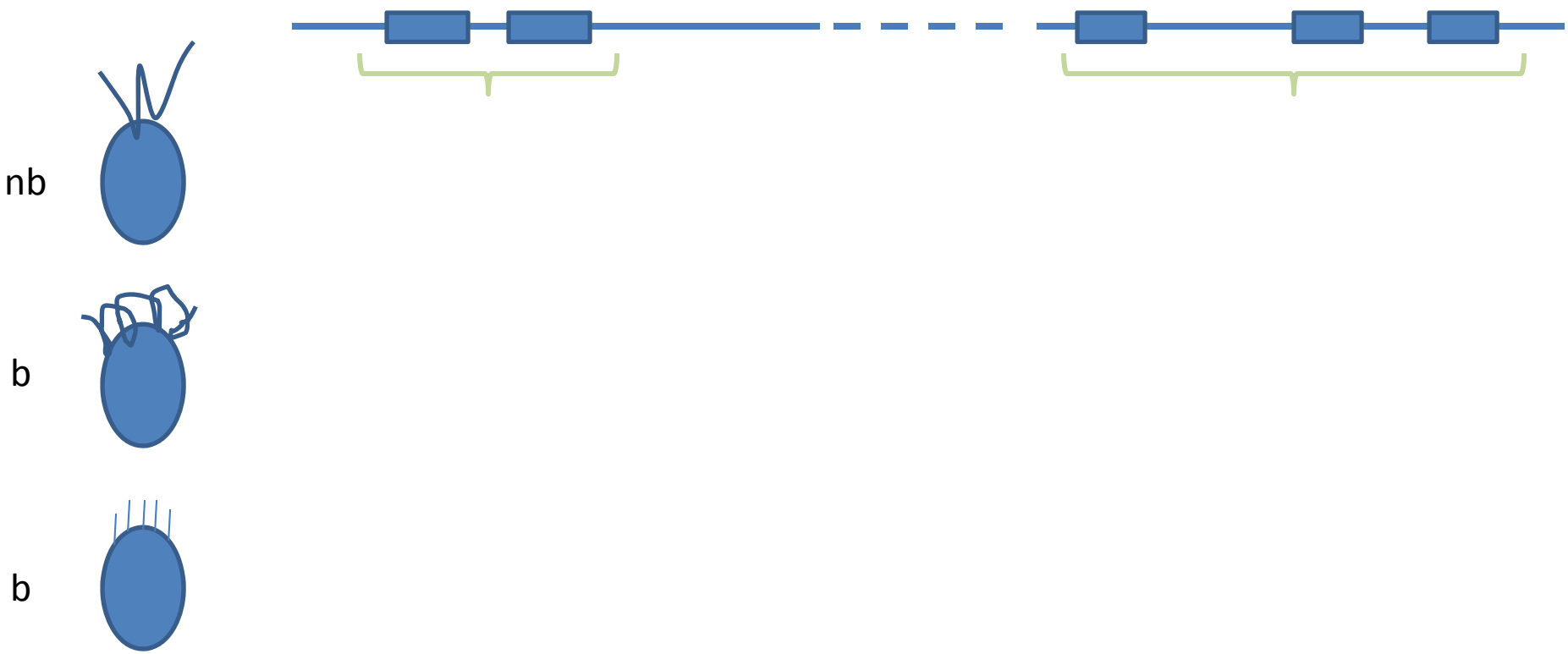
Illustrative Case Discussion - hypertension

Robert Yu

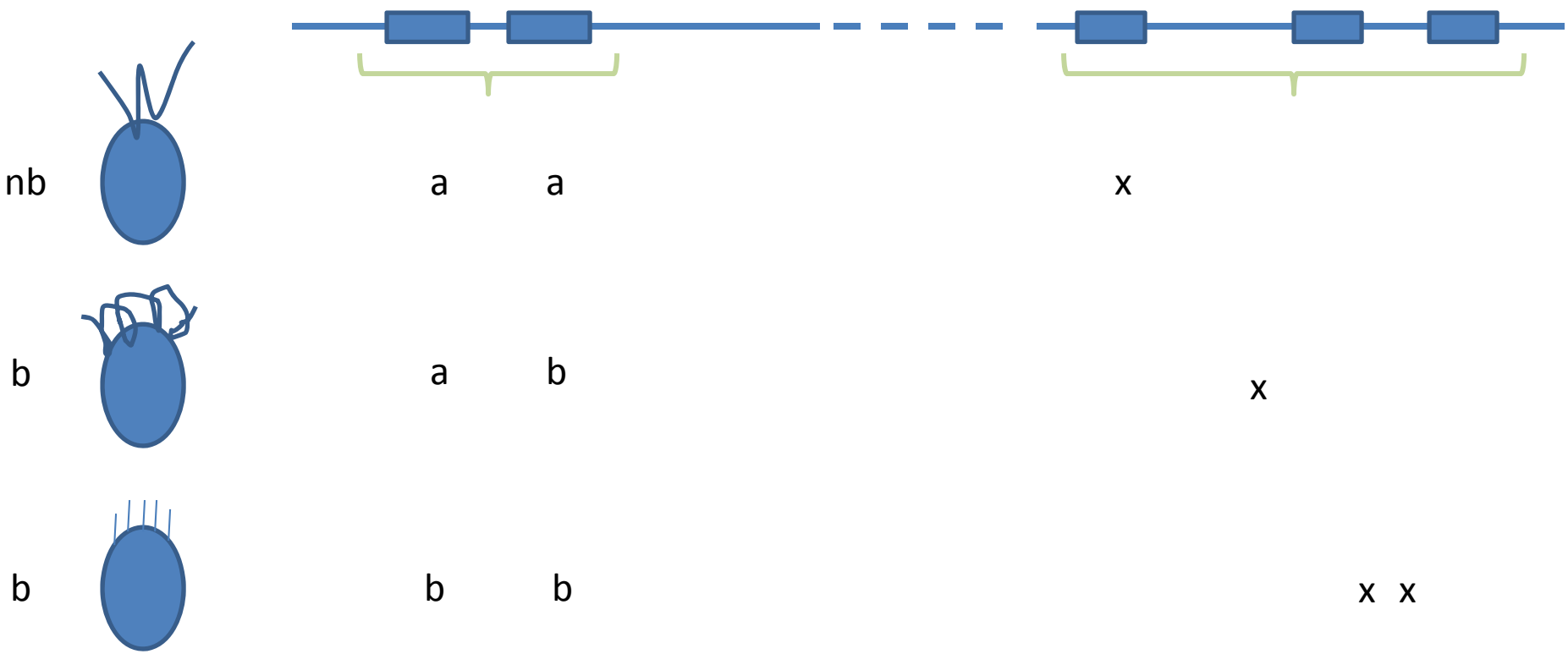
April 2015

What is “common disease, common variants”?

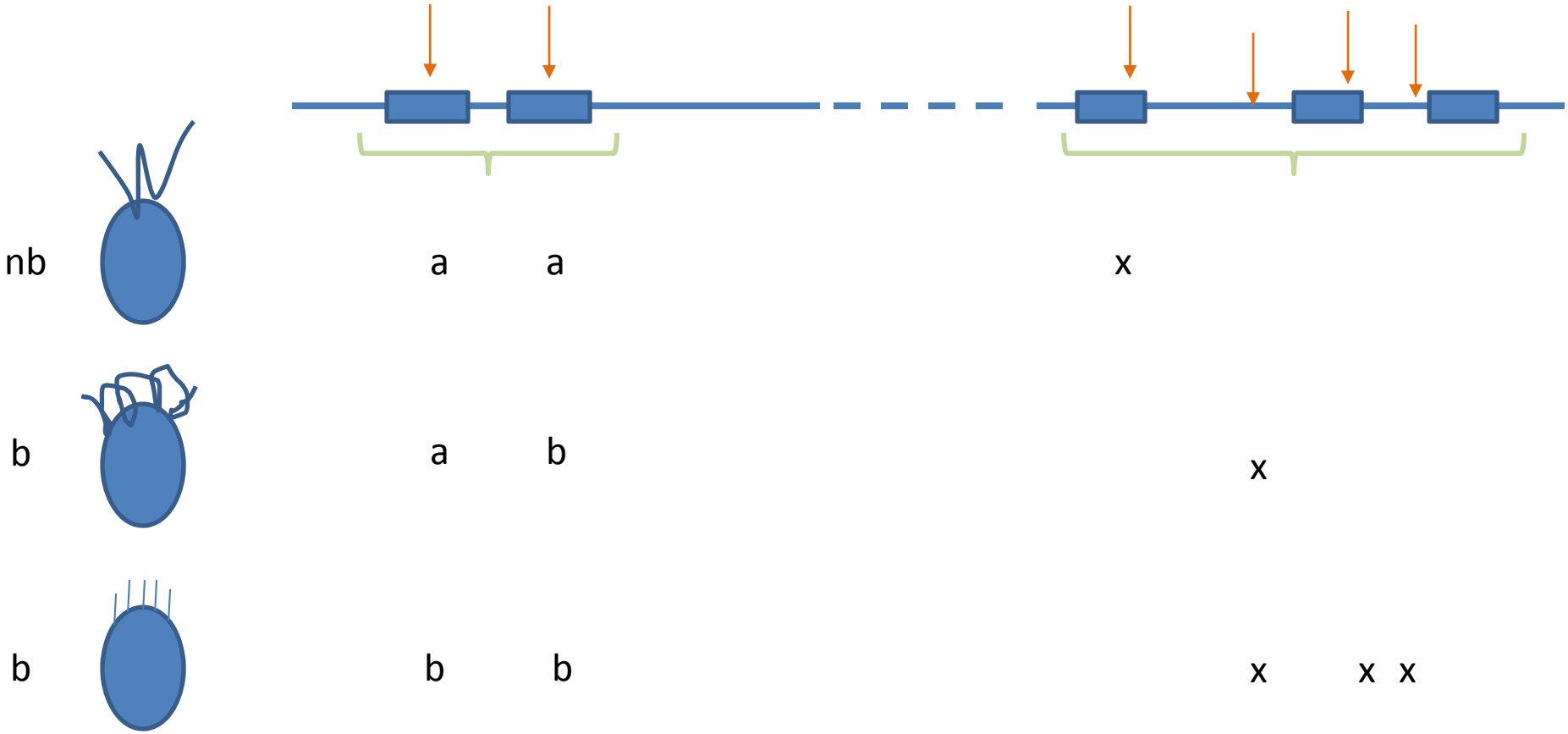
What is “common disease, common variants”?



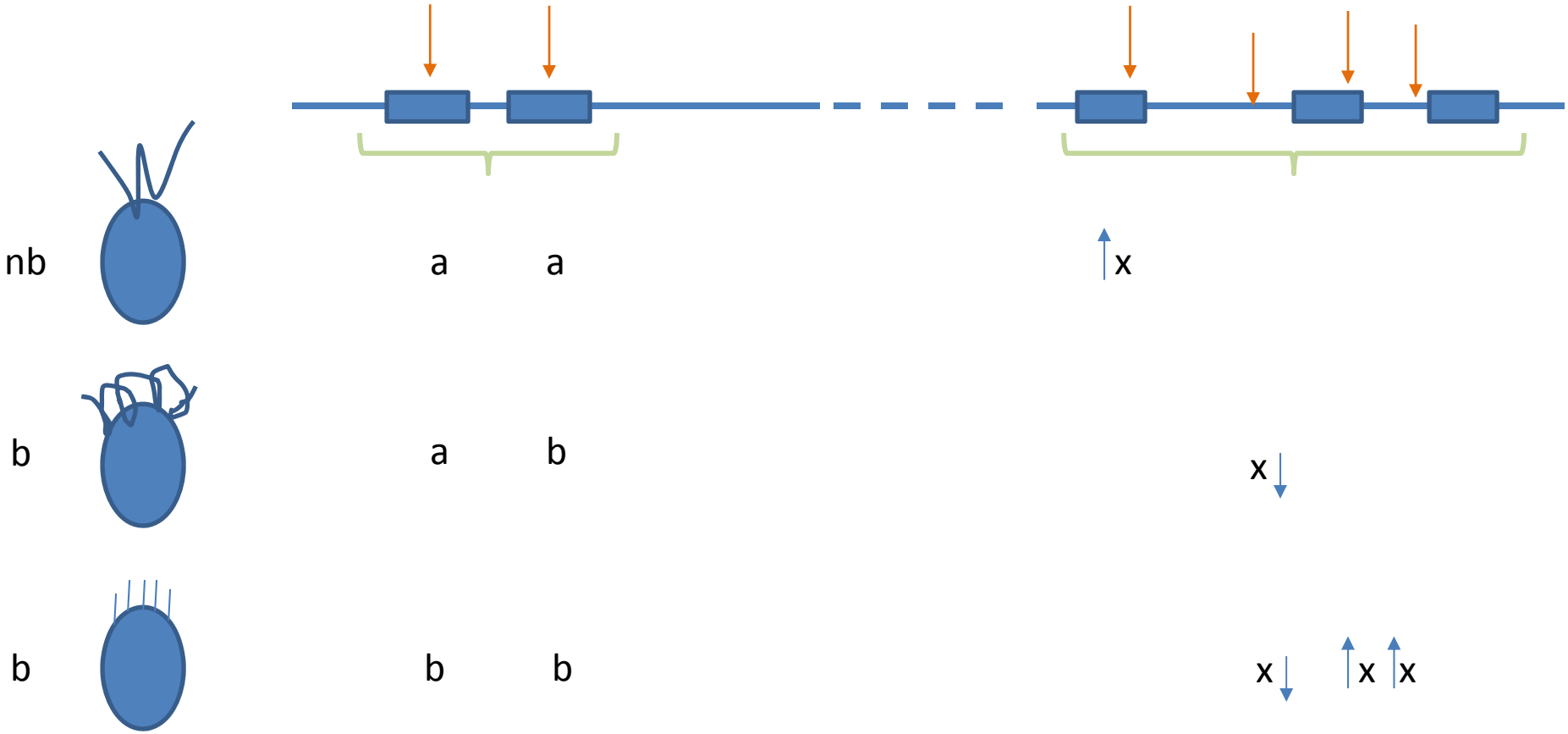
What is “common disease, common variants”?



What is “common disease, common variants”?



What is “common disease, common variants”?



Genome-wide association study of blood pressure and hypertension

2009

Daniel Levy[1,2,29], Georg B Ehret[3,4,29], Kenneth Rice[5,29], Germaine C Verwoert[6,7,28,29], Lenore J Launer[8,29], Abbas Dehghan[6], Nicole L Glazer[9], Alanna C Morrison[10], Andrew D Johnson[1,2], Thor Aspelund[11,12], Yurii Aulchenko[6], Thomas Lumley[5], Anna Köttgen[13], Ramachandran S Vasan[1,14–17], Fernando Rivadeneira[6,7], Gudny Eiriksdottir[11], Xiuqing Guo[18], Dan E Arking[3], Gary F Mitchell[19], Francesco U S Mattace-Raso[6,20], Albert V Smith[11], Kent Taylor[18], Robert B Scharpf[21], Shih-Jen Hwang[1,2], Eric J G Sijbrands[7], Joshua Bis[9], Tamara B Harris[8], Santhi K Ganesh[3,22], Christopher J O'Donnell[1,2], Albert Hofman[6], Jerome I Rotter[18], Josef Coresh[13], Emelia J Benjamin[1,14–17], André G Uitterlinden[6,7], Gerardo Heiss[23], Caroline S Fox[1,2], Jacqueline C M Witteman[6,28], Eric Boerwinkle[10], Thomas J Wang[1,24], Vilmundur Gudnason[11,12,29], Martin G Larson[1,25,29], Aravinda Chakravarti[3,13,29], Bruce M Psaty[26,27,29] & Cornelia M van Duijn[6,29]

Blood pressure is a major cardiovascular disease risk factor. To date, few variants associated with interindividual blood pressure variation have been identified and replicated. Here we report results of a genome-wide association study of systolic (SBP) and diastolic (DBP) blood pressure and hypertension in the CHARGE Consortium ($n = 29,136$), identifying 13 SNPs for SBP, 20 for DBP and 10 for hypertension at $P < 4 \times 10^{-7}$. The top ten loci for SBP and DBP were incorporated into a risk score; mean BP and prevalence of hypertension increased in relation to the number of risk alleles carried. When ten CHARGE SNPs for each trait were included in a joint meta-analysis with the Global BPgen Consortium ($n = 34,433$), four CHARGE loci attained genome-wide significance ($P < 5 \times 10^{-8}$) for SBP (*ATP2B1, CYP17A1, PLEKHA7, SH2B3*), six for DBP (*ATP2B1, CACNB2, CSK-ULK3, SH2B3, TBX3-TBX5, ULK4*) and one for hypertension (*ATP2B1*). Identifying genes associated with blood pressure advances our understanding of blood pressure regulation and highlights potential drug targets for the prevention or treatment of hypertension.

High blood pressure affects about one-third of adults and contributes to 13.5 million deaths worldwide each year and about half the global risk for stroke and ischemic heart disease[1,2]. Clinical trials, dating back more than 40 years, have proven that drug treatment to lower blood

complementary approaches have been used to search for genes associated with interindividual variation in blood pressure in the general population, but these have yielded relatively few clues. Despite considerable knowledge about pathways that are critical to blood pressure homeostasis,

Genome-wide association study of blood pressure and hypertension

Daniel Levy[1,2,29], Georg B Ehret[3,4,29], Kenneth Rice[5,29], Germaine C Verwoert[6,7,28,29], Lenore J Launer[8,29], Abbas Dehghan[6], Nicole L Glazer[9], Alanna C Morrison[10], Andrew D Johnson[1,2], Thor Aspelund[11,12], Yurii Aulchenko[6], Thomas Lumley[5], Anna Köttgen[13], Ramachandran S Vasan[1,14–17], Fernando Rivadeneira[6,7], Gudny Eiriksdottir[11], Xiuqing Guo[18], Dan E Arking[3], Gary F Mitchell[19], Francesco U S Mattace-Raso[6,20], Albert V Smith[11], Kent Taylor[18], Robert B Scharpf[21], Shih-Jen Hwang[1,2], Eric J G Sijbrands[7], Joshua Bis[9], Tamara B Harris[8], Santhi K Ganesh[3,22], Christopher J O'Donnell[1,2], Albert Hofman[6], Jerome I Rotter[18], Josef Coresh[13], Emelia J Benjamin[1,14–17], André G Uitterlinden[6,7], Gerardo Heiss[23], Caroline S Fox[1,2], Jacqueline C M Witteman[6,28], Eric Boerwinkle[10], Thomas J Wang[1,24], Vilmundur Gudnason[11,12,29], Martin G Larson[1,25,29], Aravinda Chakravarti[3,13,29], Bruce M Psaty[26,27,29] & Cornelia M van Duijn[6,29]

Blood pressure is a major cardiovascular disease risk factor. To date, few variants associated with interindividual blood pressure variation have been identified and replicated. Here we report results of a genome-wide association study of systolic (SBP) and diastolic (DBP) blood pressure and hypertension in the CHARGE Consortium ($n = 29{,}136$), identifying 13 SNPs for SBP, 20 for DBP and 10 for hypertension at $P < 4 \times 10^{-7}$. The top ten loci for SBP and DBP were incorporated into a risk score; mean BP and prevalence of hypertension increased in relation to the number of risk alleles carried. When ten CHARGE SNPs for each trait were included in a joint meta-analysis with the Global BPgen Consortium ($n = 34{,}433$), four CHARGE loci attained genome-wide significance ($P < 5 \times 10^{-8}$) for SBP (*ATP2B1, CYP17A1, PLEKHA7, SH2B3*), six for DBP (*ATP2B1, CACNB2, CSK-ULK3, SH2B3, TBX3-TBX5, ULK4*) and one for hypertension (*ATP2B1*). Identifying genes associated with blood pressure advances our understanding of blood pressure regulation and highlights potential drug targets for the prevention or treatment of hypertension.

High blood pressure affects about one-third of adults and contributes to 13.5 million deaths worldwide each year and about half the global risk for stroke and ischemic heart disease[1,2]. Clinical trials, dating back more than 40 years, have proven that drug treatment to lower blood

complementary approaches have been used to search for genes associated with interindividual variation in blood pressure in the general population, but these have yielded relatively few clues. Despite considerable knowledge about pathways that are critical to blood pressure homeostasis,

Table 2 Genome-wide association results for DBP-associated SNPs with *P* value $< 1 \times 10^{-6}$ sorted by DBP meta-analysis *P* value

					CHARGE meta-analysis, DBP			CHARGE meta-analysis, SBP			CHARGE meta-analysis, hypertension		
SNP identifier	Chr	Position	Gene	MAF	Beta	s.e.	*P*	Beta	s.e.	*P*	Beta	s.e.	*P*
rs3184504	12	110368991	*SH2B3*	0.48	**0.50**	**0.09**	**1.7E−08**	0.75	0.15	5.7E−07	0.07	0.02	7.4E−04
rs653178	12	110492139	*ATXN2*	0.48	**0.50**	**0.09**	**2.0E−08**	0.74	0.15	8.5E−07	0.07	0.02	7.7E−04
rs2681472	12	88533090	*ATP2B1*	0.17	**−0.64**	**0.12**	**3.7E−08**	−1.29	0.19	3.5E−11	−0.16	0.03	1.7E−08
rs4766578	12	110388754	*ATXN2*	0.49	**0.49**	**0.09**	**4.2E−08**	0.73	0.15	1.2E−06	0.06	0.02	1.9E−03
rs10774625	12	110394602	*ATXN2*	0.49	**0.49**	**0.09**	**4.2E−08**	0.73	0.15	1.1E−06	0.06	0.02	1.8E−03
rs2681492	12	88537220	*ATP2B1*	0.19	**−0.62**	**0.11**	**4.6E−08**	−1.26	0.18	3.0E−11	−0.14	0.03	8.4E−08
rs11105354	12	88550654	*ATP2B1*	0.17	**−0.63**	**0.12**	**5.8E−08**	−1.30	0.19	3.7E−11	−0.16	0.03	1.8E−08
rs17630235	12	111076069	*TRAFD1*	0.43	**0.50**	**0.09**	**1.0E−07**	0.69	0.15	1.1E−05	0.06	0.02	4.3E−03
rs17249754	12	88584717		0.17	**−0.63**	**0.12**	**1.0E−07**	−1.30	0.19	5.2E−11	−0.16	0.03	2.2E−08
rs11066188	12	111095097	*C12orf51*	0.43	**0.50**	**0.09**	**1.1E−07**	0.68	0.15	1.3E−05	0.06	0.02	4.2E−03
rs11105364	12	88593407		0.17	**−0.63**	**0.12**	**1.2E−07**	−1.30	0.19	4.8E−11	−0.16	0.03	2.1E−08
rs11105368	12	88598572		0.17	**−0.63**	**0.12**	**1.2E−07**	−1.30	0.19	5.3E−11	−0.16	0.03	2.2E−08
rs12579302	12	88574634		0.17	**−0.62**	**0.12**	**1.2E−07**	−1.29	0.19	6.2E−11	−0.16	0.03	2.2E−08
rs2384550	12	113837114	*TBX3-TBX5*	0.35	**−0.48**	**0.09**	**1.3E−07**	−0.71	0.15	4.3E−06	−0.08	0.02	5.6E−05
rs1991391	12	113837049		0.35	**−0.48**	**0.09**	**1.4E−07**	−0.71	0.15	3.8E−06	−0.09	0.02	5.6E−05
rs6489992	12	113837152		0.37	**−0.48**	**0.09**	**2.0E−07**	−0.71	0.15	4.7E−06	−0.08	0.02	1.9E−04
rs11065987	12	110556807		0.42	**0.48**	**0.09**	**2.2E−07**	0.70	0.15	9.4E−06	0.06	0.02	4.1E−03

Table 1 Genome-wide association results for SBP-associated SNPs with *P* value $< 1 \times 10^{-6}$ sorted by SBP meta-analysis *P* value

					CHARGE meta-analysis, SBP			CHARGE meta-analysis, DBP			CHARGE meta-analysis, hypertension		
SNP identifier	Chr	Position	Gene	MAF	Beta	s.e.	*P*	Beta	s.e.	*P*	Beta	s.e.	*P*
rs2681492	12	88537220	*ATP2B1*	0.20	**−1.26**	**0.19**	**3.0E−11**	−0.62	0.11	4.6E−08	−0.14	0.03	8.4E−08
rs2681472	12	88533090	*ATP2B1*	0.18	**−1.29**	**0.19**	**3.5E−11**	−0.64	0.11	3.7E−08	−0.16	0.03	1.7E−08
rs11105354	12	88550654	*ATP2B1*	0.18	**−1.30**	**0.20**	**3.7E−11**	−0.63	0.11	5.8E−08	−0.16	0.03	1.8E−08
rs11105364	12	88593407		0.18	**−1.30**	**0.20**	**4.8E−11**	−0.63	0.12	1.2E−07	−0.16	0.03	2.1E−08
rs17249754	12	88584717		0.18	**−1.30**	**0.20**	**5.2E−11**	−0.63	0.12	1.0E−07	−0.16	0.03	2.2E−08
rs11105368	12	88598572		0.18	**−1.30**	**0.20**	**5.3E−11**	−0.63	0.12	1.3E−07	−0.16	0.03	2.2E−08
rs12579302	12	88574634		0.18	**−1.29**	**0.20**	**6.2E−11**	−0.62	0.12	1.3E−07	−0.16	0.03	2.2E−08
rs12230074	12	88614998		0.17	**−1.31**	**0.20**	**9.1E−11**	−0.62	0.12	3.4E−07	−0.17	0.03	2.9E−08
rs11105378	12	88614872		0.17	**−1.31**	**0.20**	**9.1E−11**	−0.62	0.12	3.1E−07	−0.17	0.03	2.8E−08
rs4842666	12	88465680		0.17	**−1.20**	**0.21**	**6.5E−09**	−0.62	0.12	4.5E−07	−0.15	0.03	3.4E−07
rs8096897	18	13428905	*C18orf1*	0.01	**−12.87**	**2.33**	**3.2E−08**	−4.07	1.33	2.9E−03	−0.73	0.35	0.04
rs11105328	12	88466521		0.18	**−1.11**	**0.20**	**4.2E−08**	−0.61	0.12	5.1E−07	−0.15	0.03	7.1E−07
rs880315	1	10719453	*CASZ1*	0.35	**0.89**	**0.17**	**2.1E−07**	0.30	0.10	2.9E−03	0.09	0.02	6.2E−05
rs3184504	12	110368991	*SH2B3*	0.48	**0.75**	**0.15**	**5.7E−07**	0.50	0.09	1.7E−08	0.07	0.02	7.4E−04
rs381815	11	16858844	*PLEKHA7*	0.26	**0.84**	**0.17**	**5.8E−07**	0.51	0.10	4.3E−07	0.09	0.02	1.7E−04
rs7926335	11	16874445	*PLEKHA7*	0.26	**0.85**	**0.17**	**5.8E−07**	0.51	0.10	4.8E−07	0.09	0.02	1.9E−04
rs7571613	2	190513907	*PMS1*	0.18	**0.96**	**0.19**	**7.2E−07**	0.55	0.11	2.2E−06	0.09	0.03	5.2E−04
rs11895934	2	190510498		0.18	**0.96**	**0.19**	**7.3E−07**	0.55	0.11	2.2E−06	0.09	0.03	5.5E−04
rs7564968	2	190520217		0.18	**0.96**	**0.19**	**8.0E−07**	0.55	0.11	2.3E−06	0.09	0.03	4.9E−04
rs653178	12	110492139	*ATXN2*	0.48	**0.74**	**0.15**	**8.5E−07**	0.50	0.09	2.0E−08	0.07	0.02	7.8E−04
rs284277	1	10713384	*CASZ1*	0.35	**0.79**	**0.16**	**9.4E−07**	0.24	0.09	0.01	0.09	0.02	6.9E−05

MAF, minor allele frequency. Beta is the effect size on blood pressure in mm Hg, per allele based on the additive genetic model.

On the Aug 2010 annotation of the human genome

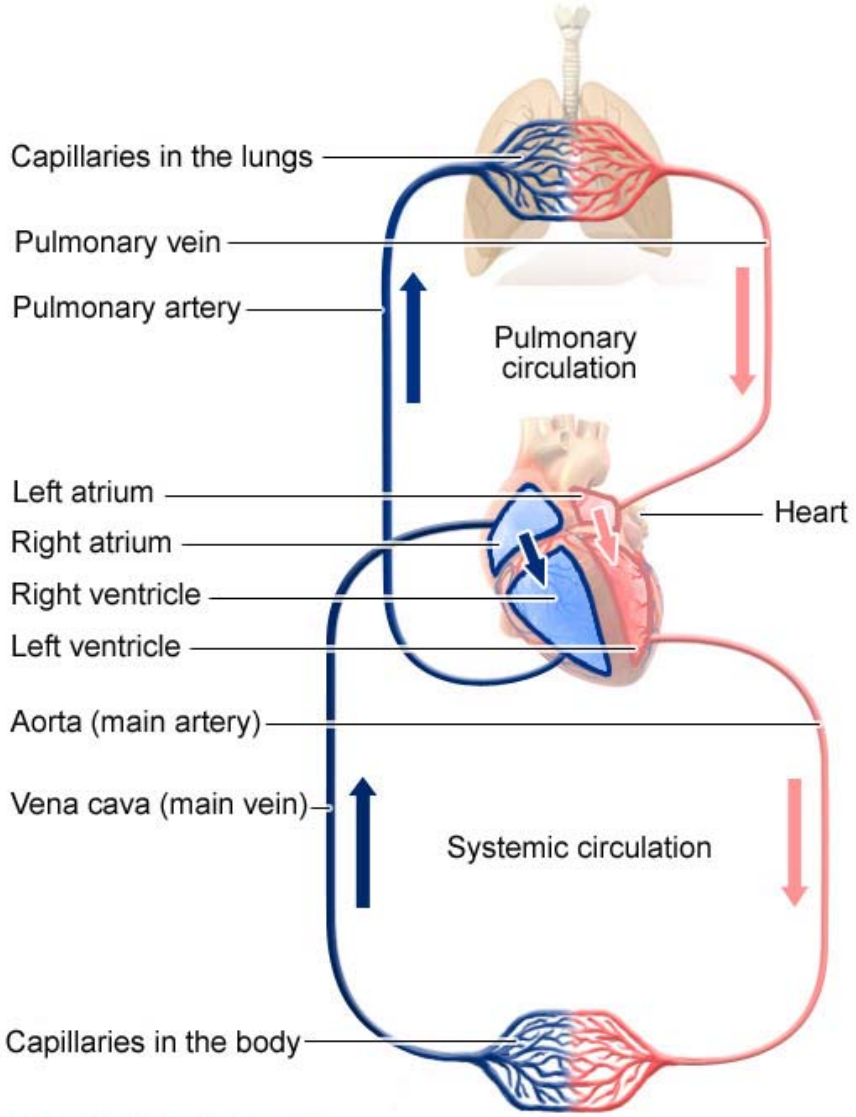
NCBI's database

770 genes relate directly or indirectly to Hypertension

Page 1 of 8 go to next last page.

	Gene Name	Aligned on chrom	Cyto location	Supporting cDNA clones	Description
1	ABCA1	9	"9q31.1"	222	ATP-binding cassette, sub-family A (ABC1), member 1.
2	ABCB1	7	"7q21.12"	153	ATP-binding cassette, sub-family B (MDR/TAP), member 1.
3	ABCB8andACCN3	7	"7q36"	567	amiloride-sensitive cation channel 3 and ATP -binding cassette, sub-family B (MDR/TAP), member 8.
4	ABCC1	16	"16p13.1"		
5	ABCC9	12	"12p12.1"		
6	ACADSB	10	"10q26.13"		
7	ACE2	X	"Xp22"		
8	ACE3	17	"17q23.3"		
9	ACOT7	1	"1p36"		
10	ACSM1	16	"16p12.3"		
11	ACSM3	16	"16p13.11"		
12	ACSS1	20	"20p11.23-p11.		
13	ACVRL1	12	"12q11-q14"		
14	ACYP2	2	"2p16.2"		
15	ADA	20	"20q13.12"		
16	ADAMTS16	5	"5p15"		
17	ADAMTSL5	19	"19p13.3"	36	ADAMTS-like 5.
18	ADCY5	3	"3q13.2-q21"	147	adenylate cyclase 5.
755	WNK2	9	"9q22.3"	183	WNK lysine deficient protein kinase 2.
756	WNK3	X	"Xp11.22"	78	WNK lysine deficient protein kinase 3.
757	WNK4	17	"17q21-q22"	86	WNK lysine deficient protein kinase 4.
758	WNT3	17	"17q21"	62	wingless-type MMTV integration site family, member 3.
759	XCL1	1	"1q23"	29	chemokine (C motif) ligand 1.
760	XDH	2	"2p23.1"	50	xanthine dehydrogenase.
761	XYLT1	16	"16p12.3"	120	xylosyltransferase I.
762	XYLT2	17	"17q21.33"	169	xylosyltransferase II.
763	YEATS4	12	"12q13-q15"	154	YEATS domain-containing protein 4.
764	ZC3H3	8	"8q24.3"	129	zinc finger CCCH-type containing 3.
765	ZC3H11A	1	"1q32.1"	371	zinc finger CCCH-type containing 11A.
766	ZNF79	9	"9q34"	66	zinc finger protein 79.
767	ZNF607andZNF781andZFP30	19	"19q13.1"	186	zinc finger protein 30 homolog (mouse) and zinc finger protein 607 and zinc finger protein 781.
768	ZNF618	9	"9q32"	215	zinc finger protein 618.
769	ZP4	1	"1q43"	11	zona pellucida glycoprotein 4.
770	ZSWIM2	2	"2q32.1"	35	zinc finger, SWIM-type containing 2.

Exemplary Case: Hypertension



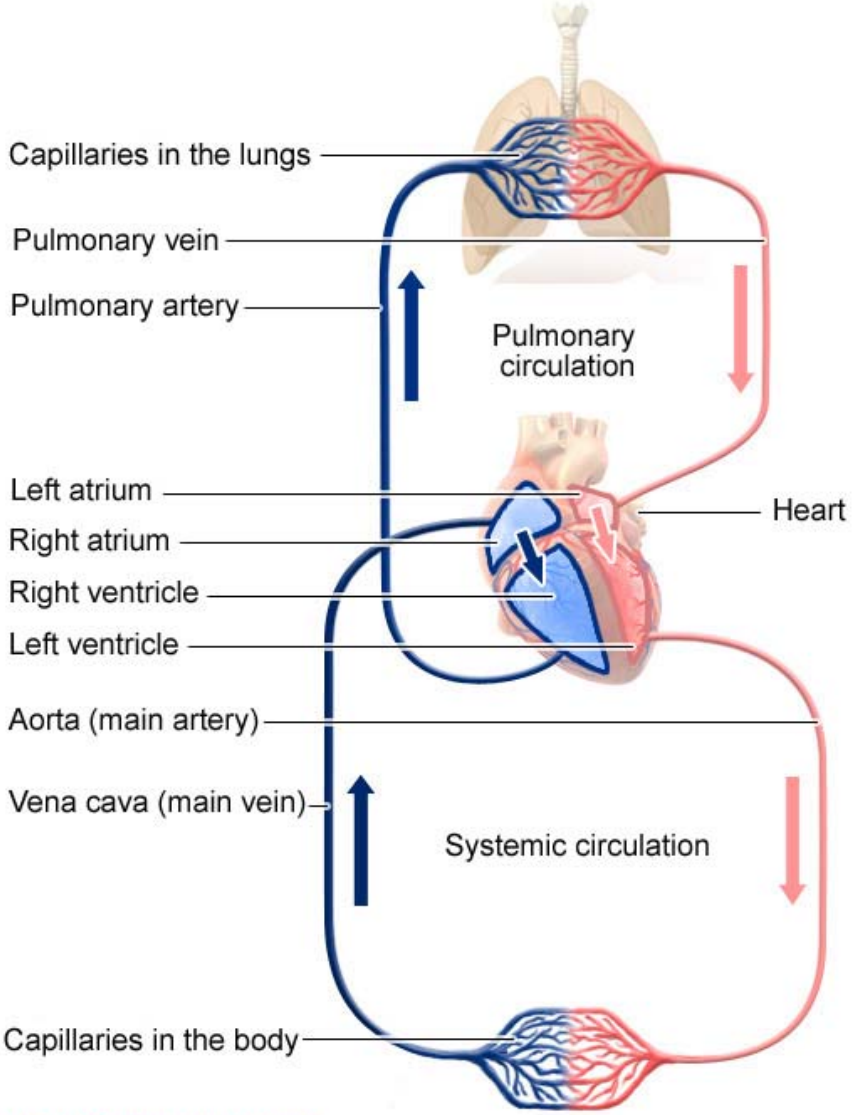
High blood pressure

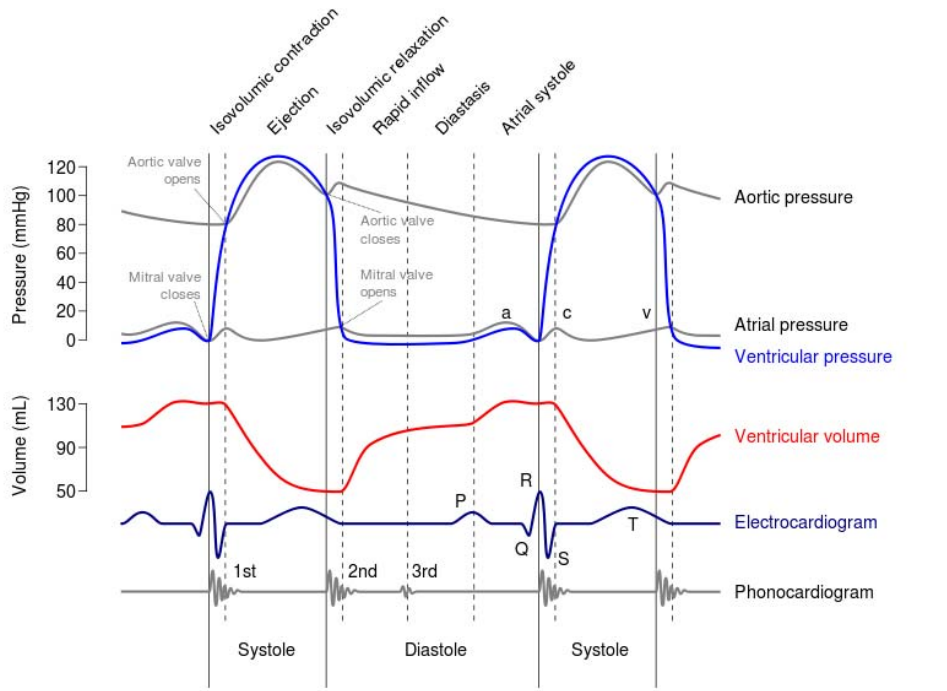
A big health risk factor and complex disease.

This has caught long time and wide range of scientific research efforts to study the causes of this disease. GWAS and NGS studies are part of the effort to find genetic elements.

Ref: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0023062/?figure=1

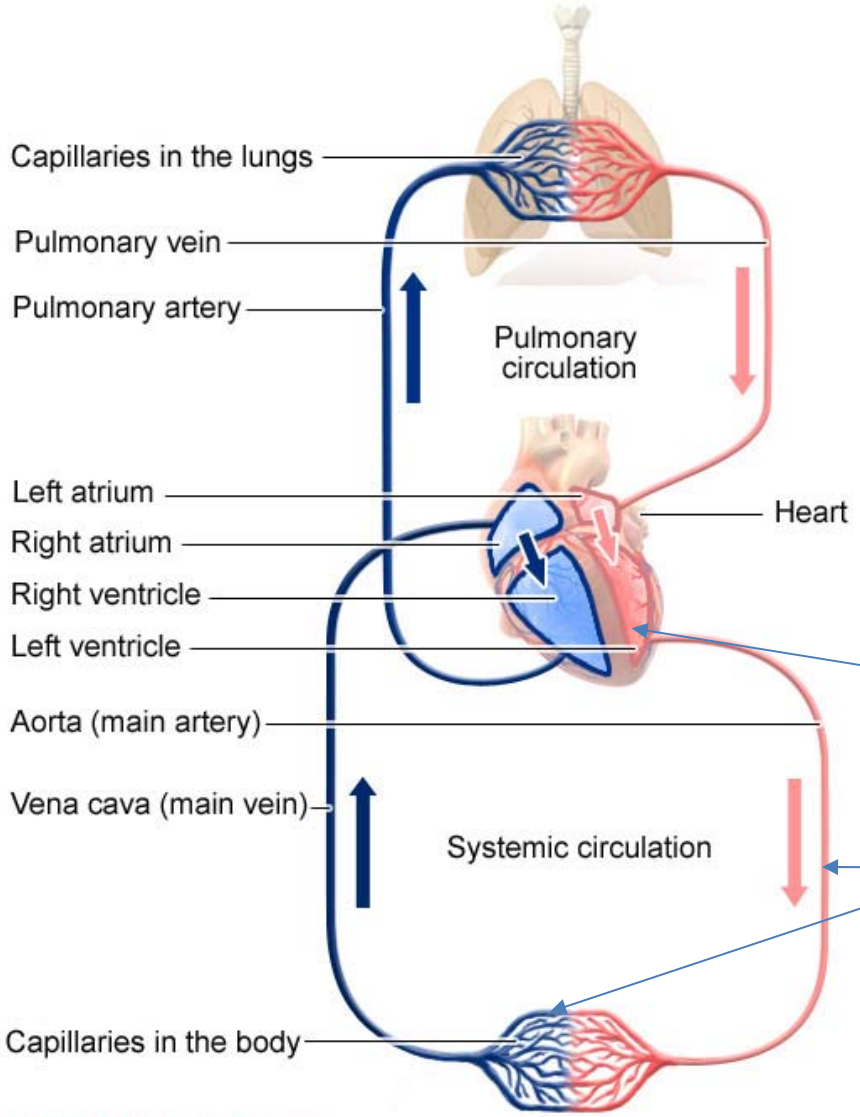
Exemplary Case: Hypertension





Ref: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0023062/?figure=1

Major Types of High Blood Pressure



Blood pressure is a phenotypic measurement
But the causal factors of the high blood pressure are many and vary from individual to individual.

Pumping too much?

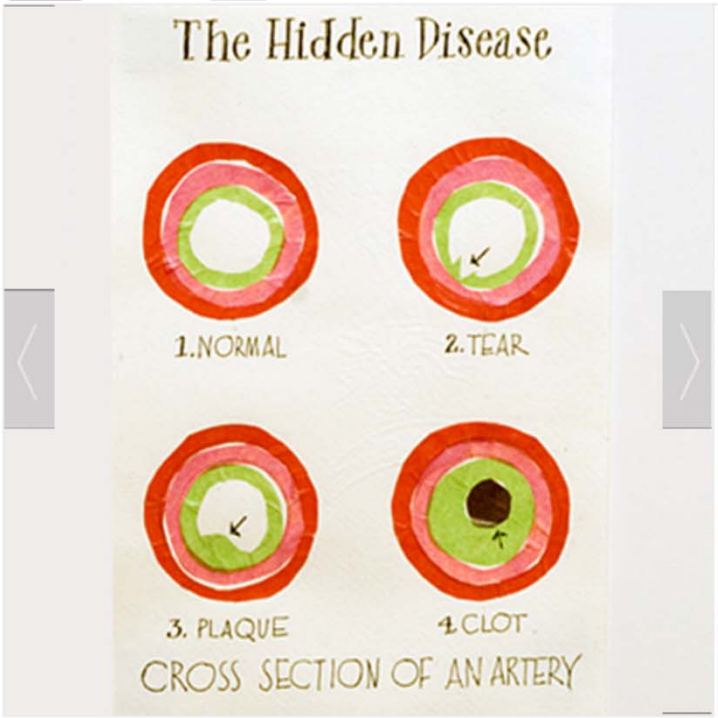
Or having higher resistance?

Various reasons and causes will lead to a higher resistance in the blood vessel.

Ref: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0023062/?figure=1

Great Number of Factors Leading to Hypertension

What Causes Clogged Arteries? A Visual Guide



2 OF 8

The hidden disease

1. A normal artery is like a new rubber band: flexible, strong, and elastic.

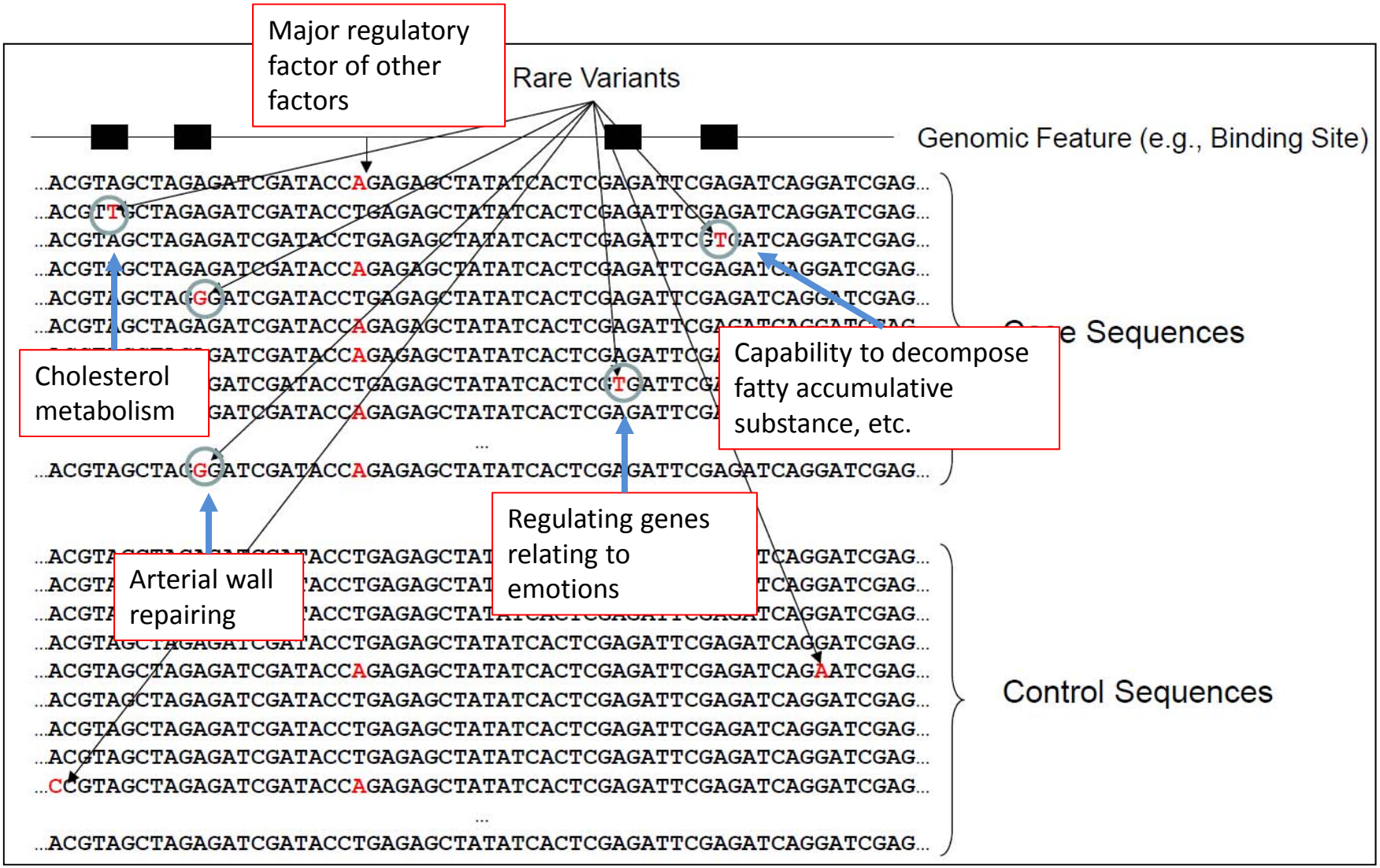
2. Although the exact trigger of atherosclerosis is unknown, researchers suspect that the process begins with damage to the inner wall of the artery (which can be caused by high cholesterol, hypertension, or cigarette smoking, among others).

3. Over time, cholesterol, calcium, and other substances accumulate in the wall of the artery and form fatty deposits called plaques. The narrower artery opening limits blood flow.

4. These plaques can burst, causing a blood clot to form.

Ref: http://www.health.com/health/gallery/0,,20307285_2,00.html

An Illustrative Picture for Hypertension Genetics?



Summaries

- Rare variants contribution
  - Hypertension in different individuals may not be having the same genetic characteristics
- Methods for rare variant analyses are important.
  - RV grouping and categorizing affect the analysis.
- Composition of samples in a data set could play a more crucial role in reaching the final results/conclusion.
  - If a set of samples are sharing more causal RVs, the results/conclusion may not be the same when the sample composition is different.
  - Simple requirement of replication may not be easy to achieve.