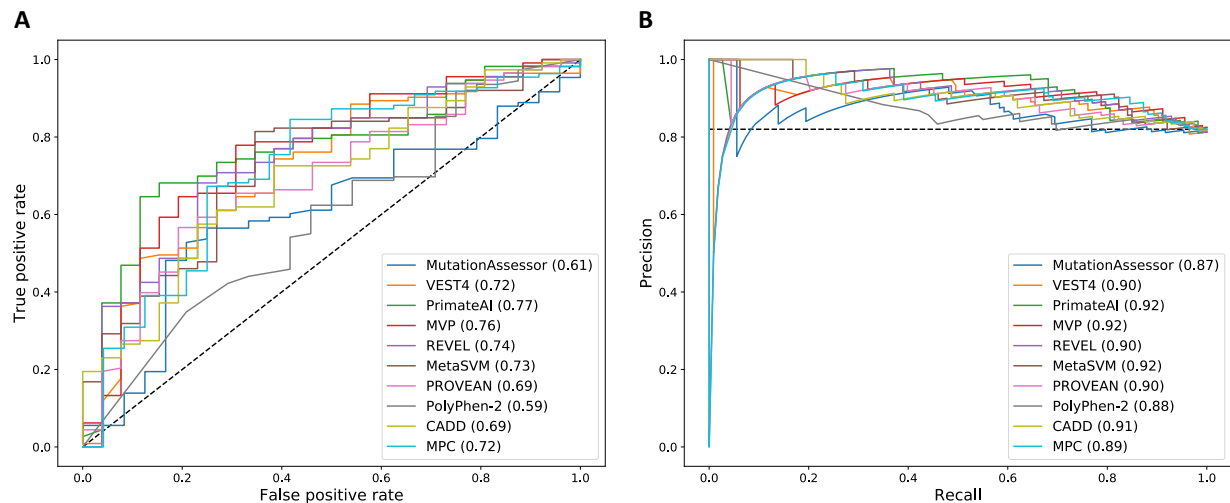
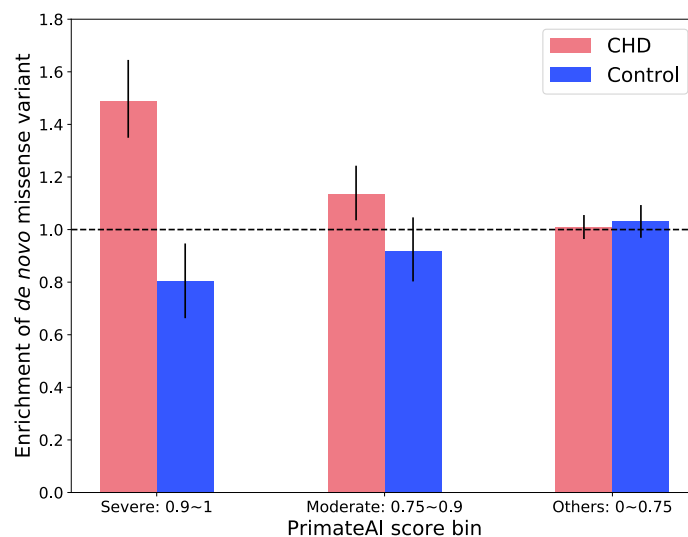
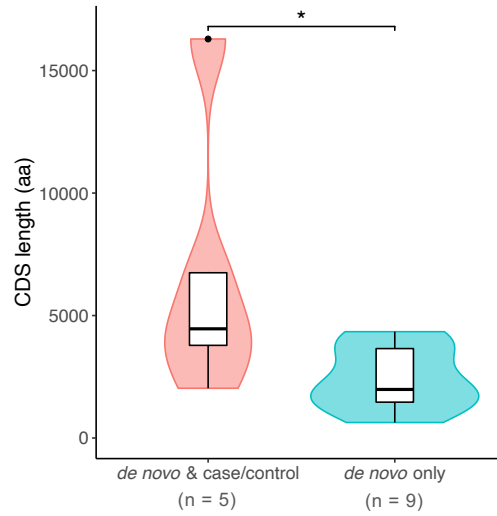




Supplementary Figure 1. Performance metrics of missense variant prediction methods. (A) Receiver operating characteristic curve for identifying CHD proband *de novo* variants. The values in parentheses are the area under the curve for each method. **(B)** Precision recall curve for identifying CHD proband *de novo* variants. The values in parentheses are the area under the curve for each method.



Supplementary Figure 2. Enrichment of *de novo* missense variants for PrimateAI score ranges. The enrichment is derived with respect to the expected number of *de novo* missense variants in cohorts of the same size as CHD trios ($n = 3,835$) and control trios ($n = 2,179$). The error bars are 95 % bootstrap confidence intervals.



Supplementary Figure 3. Coding sequence length of significant OFC genes by discovery model. Distribution of coding sequence length for the significant genes unique to the “*de novo* & case/control” model and “*de novo* only” model. The number of genes is labeled below each category. CDS, coding sequence; aa, amino acid. *: $p < 0.05$, one-sided Wilcoxon rank-sum test.

Supplementary Data 1. Number of trios for CHD, OFC, and unaffected children used for *de novo* variants.

Supplementary Data 2. List of 225 Human CHD genes.

Supplementary Data 3. TADA results for CHD

Supplementary Data 4. TADA results for OFC

Supplementary Data 5. List of 148 Human OFC genes.

Supplementary Data 6. *De novo* TF DBD missense variants in the CHD cohorts

Supplementary Data 7. *De novo* missense variants in OFC genes in OFC cohorts

Supplementary Data 8. TF DBD Pfam domains

Supplementary Data 9. *De novo* TF DBD missense variants in the CHD and OFC cohorts

PRISMA 2020 Main Checklist

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	"Meta-analysis" in the title
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	5th paragraph in the Introduction
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	5th paragraph in the Introduction
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods

Topic	No.	Item	Location where item is reported
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	"Genetic data from family trio cohorts of CHD and OFC" in the Methods
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	NA
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	NA
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	NA
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	NA
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	NA
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	NA
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	NA
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	NA
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	NA
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	NA

Topic	No.	Item	Location where item is reported
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	NA
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	"Genetic variants identified from multiple family trio cohorts of CHD and OFC" in the Results
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	NA
Study characteristics	17	Cite each included study and present its characteristics.	"Genetic variants identified from multiple family trio cohorts of CHD and OFC" in the Results
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	NA
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	NA
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	NA
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	NA
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	NA
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	NA
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	NA

Topic	No.	Item	Location where item is reported
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	NA
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	First paragraph of the Discussion
	23b	Discuss any limitations of the evidence included in the review.	Last paragraph of the Discussion
	23c	Discuss any limitations of the review processes used.	Last paragraph of the Discussion
	23d	Discuss implications of the results for practice, policy, and future research.	Last paragraph of the Discussion
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	NA
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	NA
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	NA
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Competing Interest section
Competing interests	26	Declare any competing interests of review authors.	Competing Interest section
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Code availability section

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. MetaArXiv. 2020, September 14. DOI: 10.31222/osf.io/v7gm2. For more information, visit: www.prisma-statement.org

PRIMSA Abstract Checklist

Topic	No.	Item	Reported?
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesize results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	Yes
Registration	12	Provide the register name and registration number.	No

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. MetaArXiv. 2020, September 14. DOI: 10.31222/osf.io/v7gm2. For more information, visit: www.prisma-statement.org