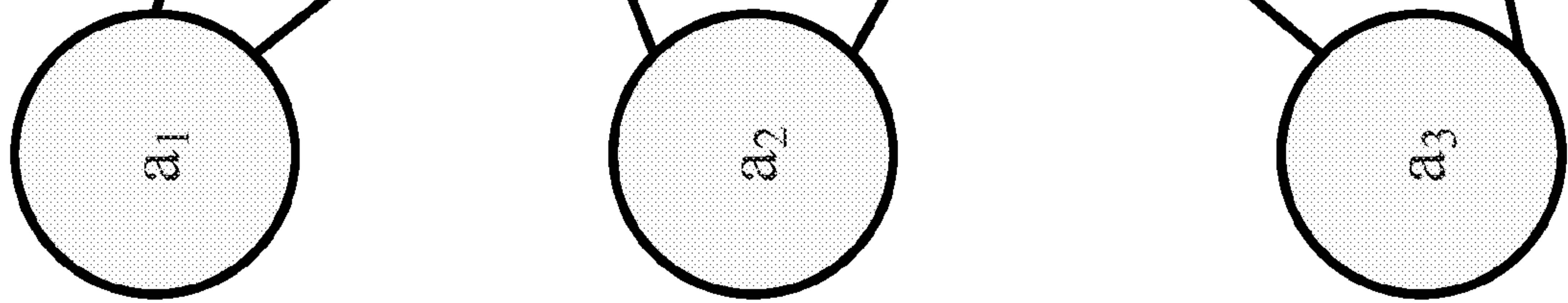


Feed-Forward Neural Network

Input



Hidden



Output



w_{11}

w_{21}

w_{12}

w_{31}

w_{22}

w_{32}

$f(\bullet)$

v_{11}

v_{12}

v_{21}

v_{22}

$f(\bullet)$

$g(\bullet)$

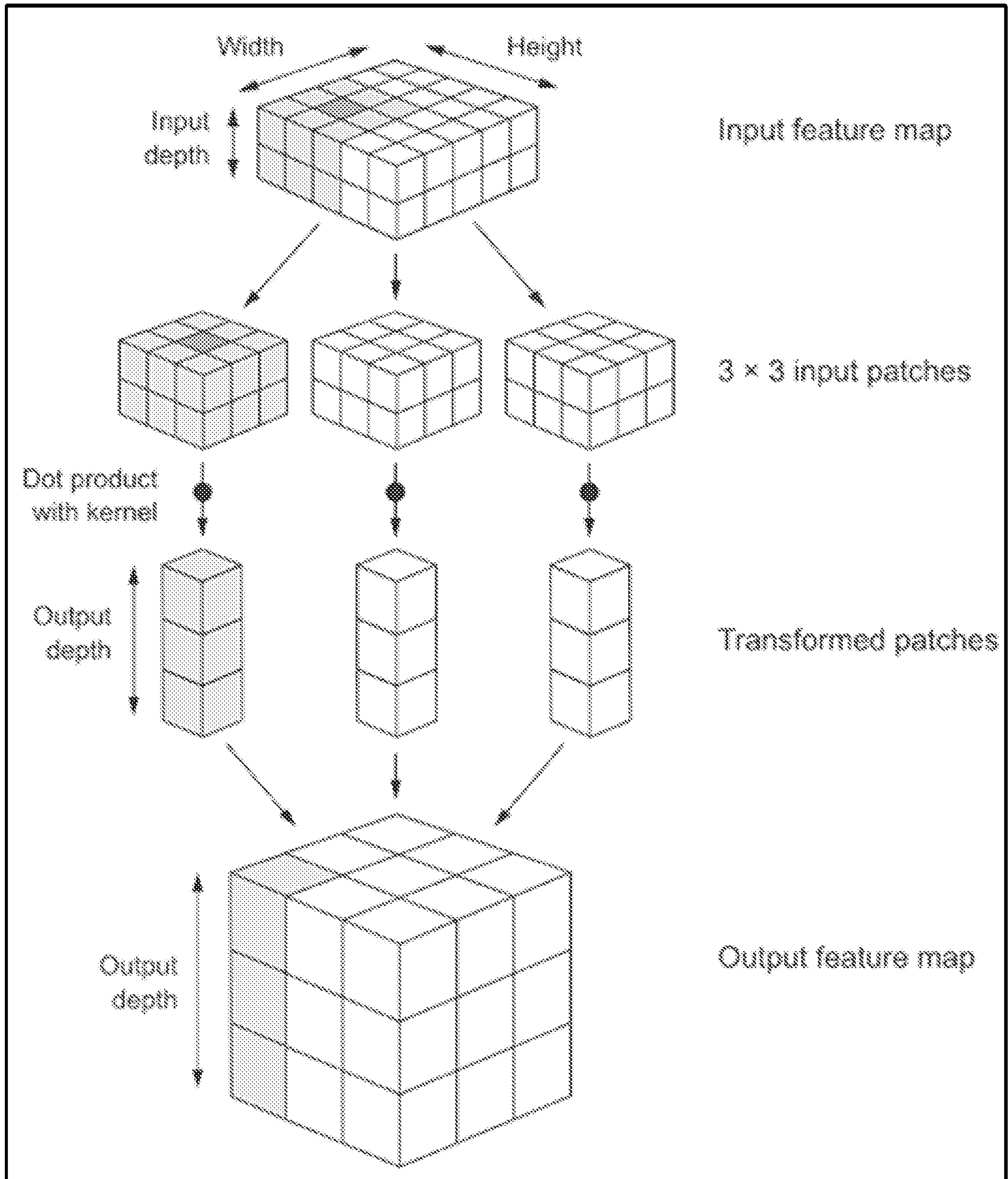
$g(\bullet)$

φ_k

φ_k

FIG. 1A

Convolutional Neural Network

**FIG. 1B**

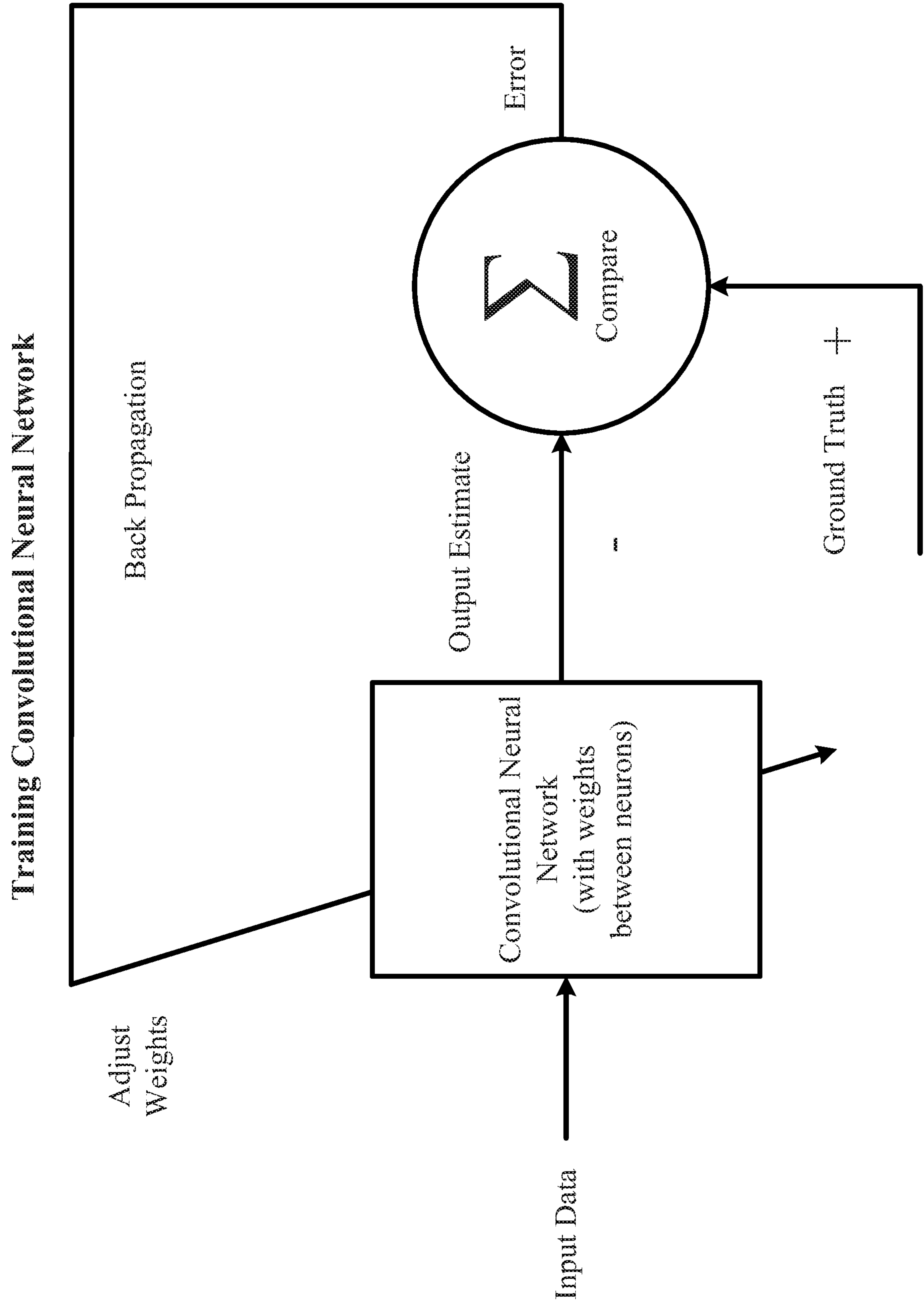


FIG. 1C

Sub-Sampling Layers of Convolutional Neural Networks

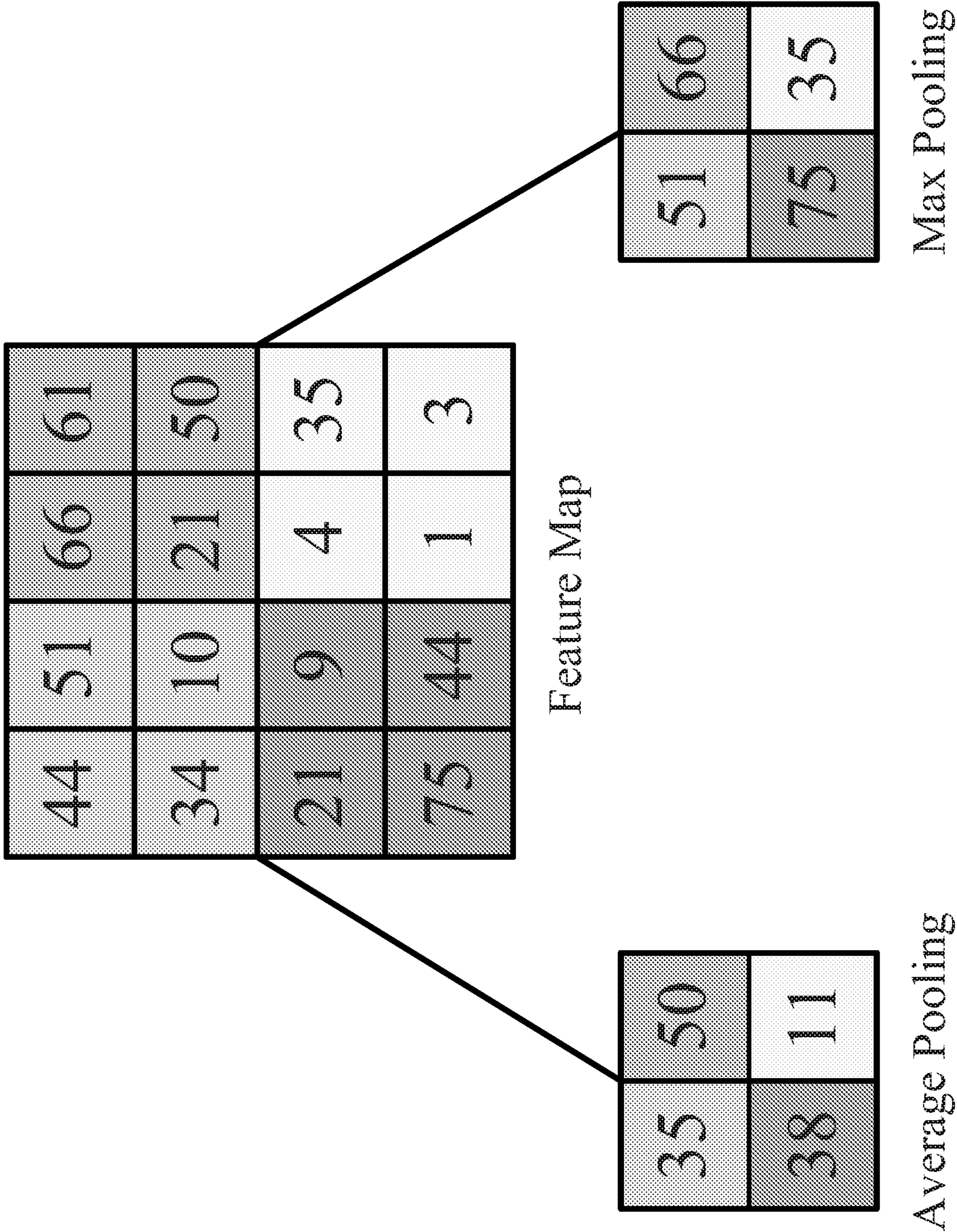
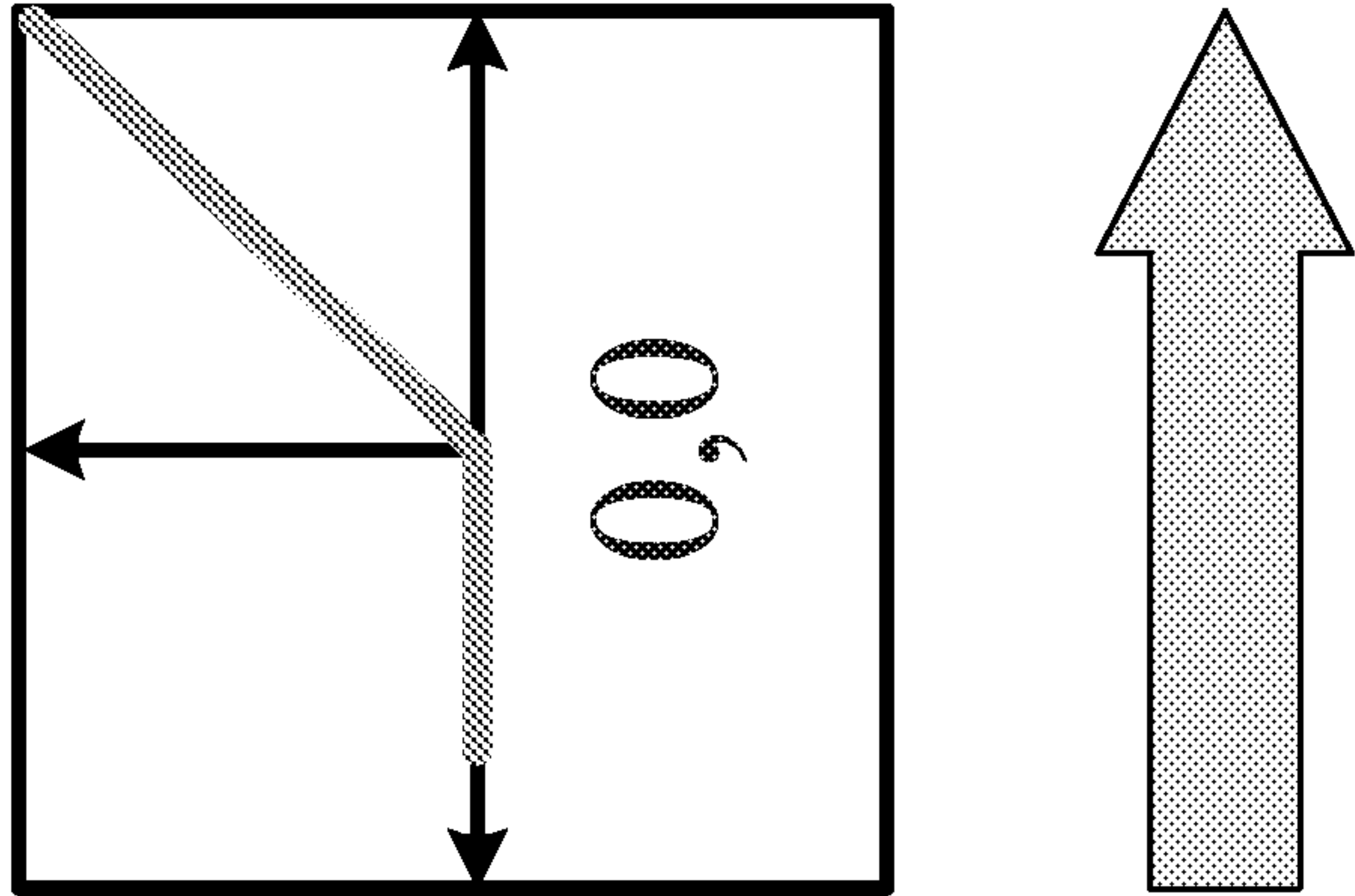


FIG. 1D

Non-Linear Layers of Convolutional Neural Networks

100	-43	76	-21
-90	-10	40	21
-19	44	1	-35
221	64	41	23



100	0	76	0
0	0	40	21
0	44	1	0
221	64	41	23

ReLU Activation Function

FIG. 1E

Convolution Layers of Convolutional Neural Networks

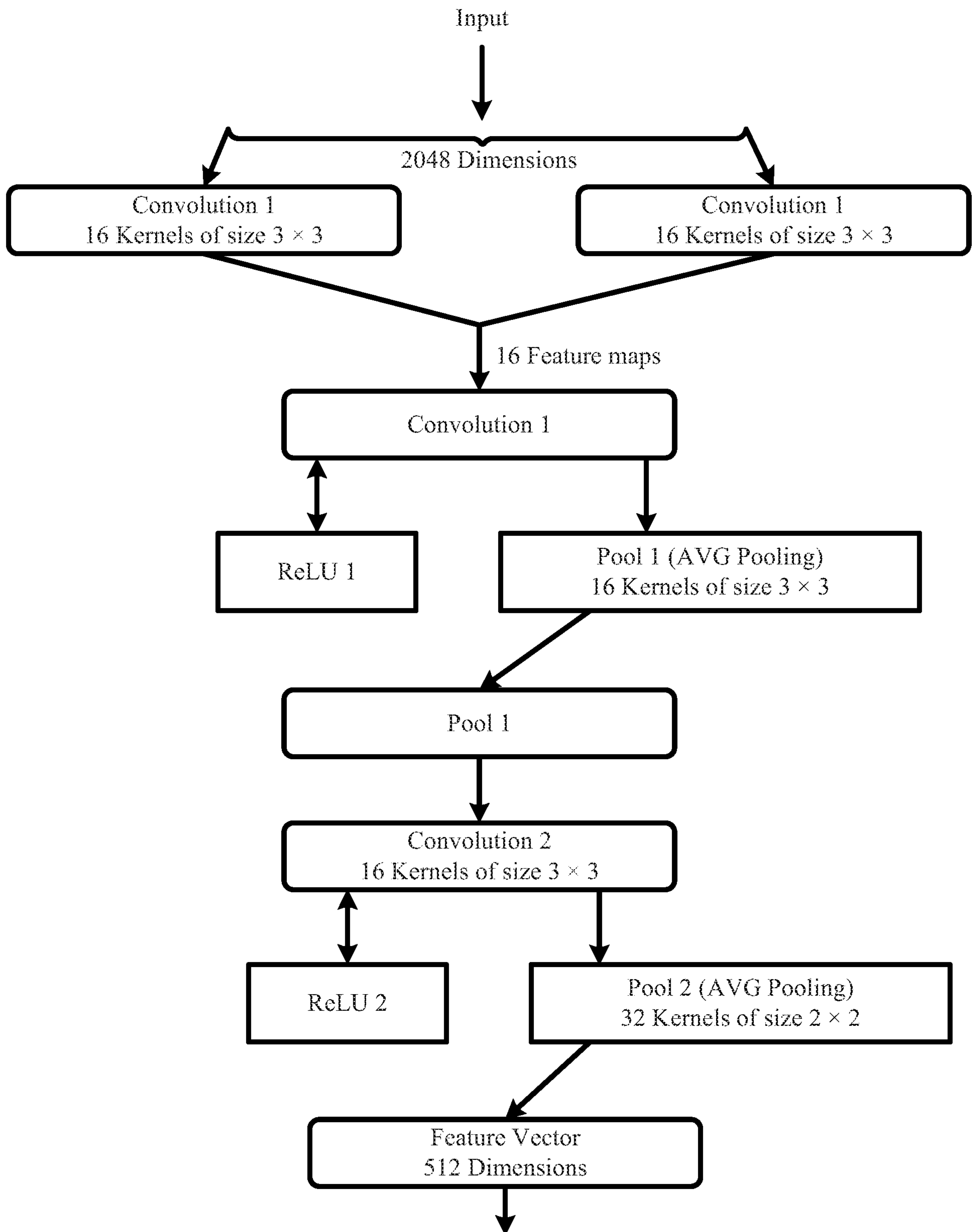


FIG. 1F

Residual Connections used in Convolutional Neural Networks

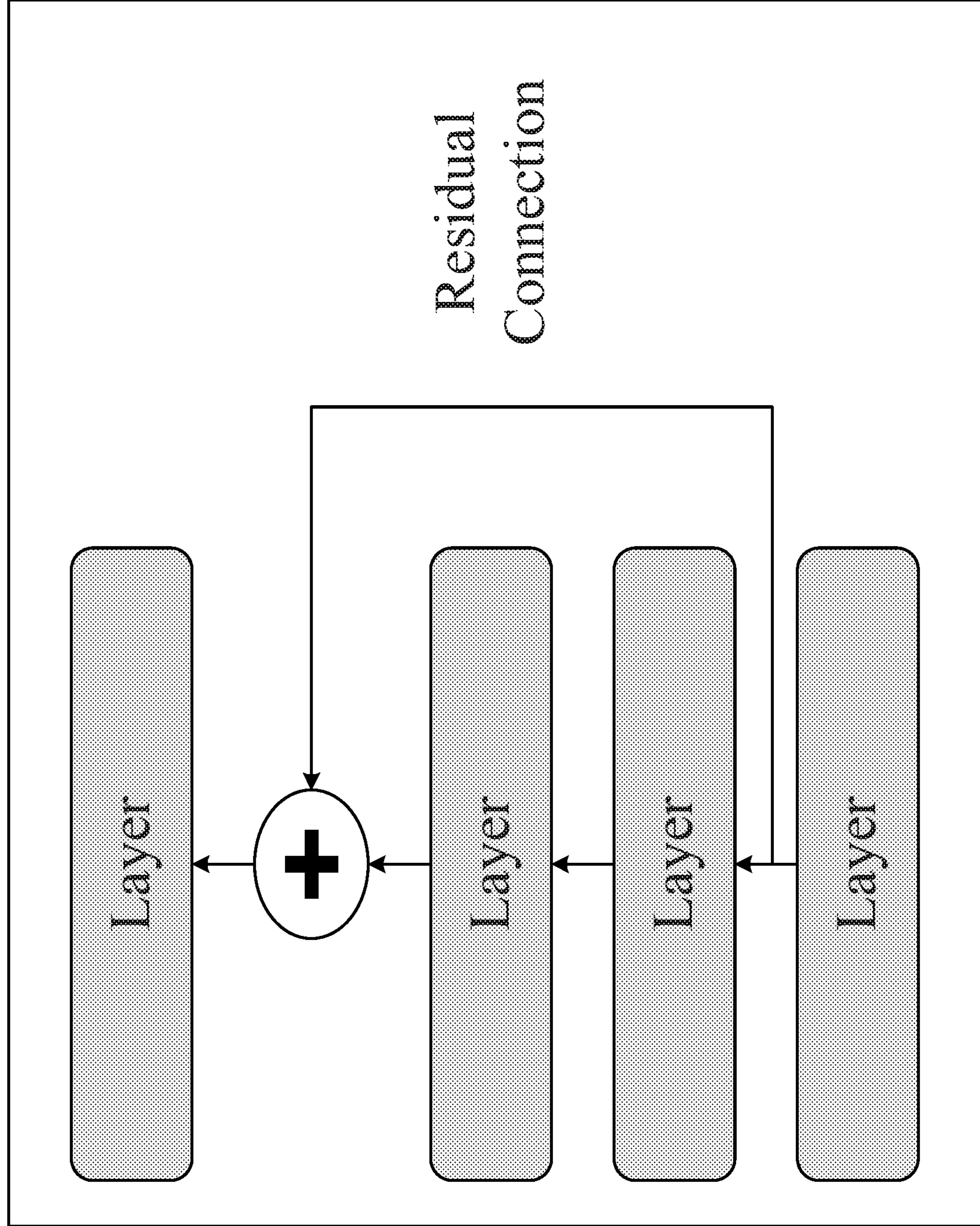
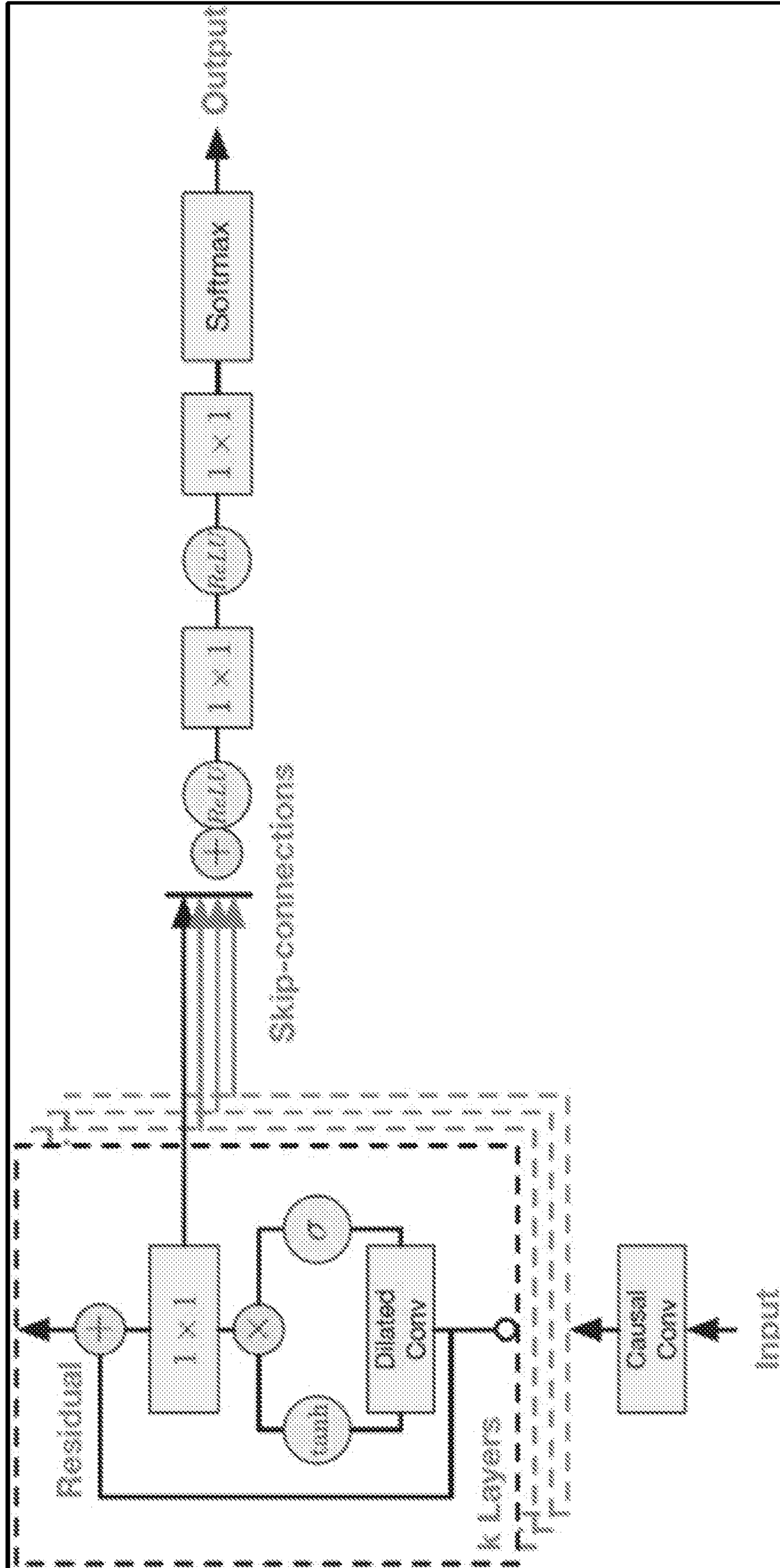


FIG. 1G



Residual Block and Skip-Connections used in
Convolutional Neural Networks

FIG. 1H

Batch Normalization Forward Pass with Convolutional Neural Networks

$$\begin{aligned}\mu_{\mathcal{B}} &= \frac{1}{n} \sum_{i=1}^n x_i^{(\ell-1)} \\ \sigma_{\mathcal{B}}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i^{(\ell-1)} - \mu_{\mathcal{B}})^2 \\ \hat{x}^{(\ell-1)} &= \frac{x^{(\ell-1)} - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \\ x^{(\ell)} &= \gamma^{(\ell)} \hat{x}^{(\ell-1)} + \beta^{(\ell)}\end{aligned}$$

FIG. 11

Batch Normalization – Inference with Convolutional Neural Networks

$$\hat{x}^{(\ell-1)} = \frac{x^{(\ell-1)} - \mu_{\mathcal{D}}}{\sqrt{\sigma_{\mathcal{D}}^2 + \epsilon}}$$
$$x_i^{(\ell)} = \gamma^{(\ell)} \hat{x}_i^{(\ell-1)} + \beta^{(\ell)}$$

FIG. 1J

Batch Normalization Backward Pass with
Convolutional Neural Networks

$$\begin{aligned}\nabla_{\gamma^{(\ell)}} \mathcal{L} &= \sum_{i=1}^n (\nabla_{x^{(\ell+1)}} \mathcal{L})_i \cdot \hat{x}_i^{(\ell)} \\ \nabla_{\beta^{(\ell)}} \mathcal{L} &= \sum_{i=1}^n (\nabla_{x^{(\ell+1)}} \mathcal{L})_i\end{aligned}$$

FIG. 1K

Batch Normalization in Convolution Layers

```
conv_model.add(layers.Conv2D(32, 3, activation='relu'))  <----- After a Conv layer
conv_model.add(layers.BatchNormalization())
dense_model.add(layers.Dense(32, activation='relu'))  <----- After a Dense layer
dense_model.add(layers.BatchNormalization())
```

FIG. 1L

1D Convolution used in Convolutional Neural Networks

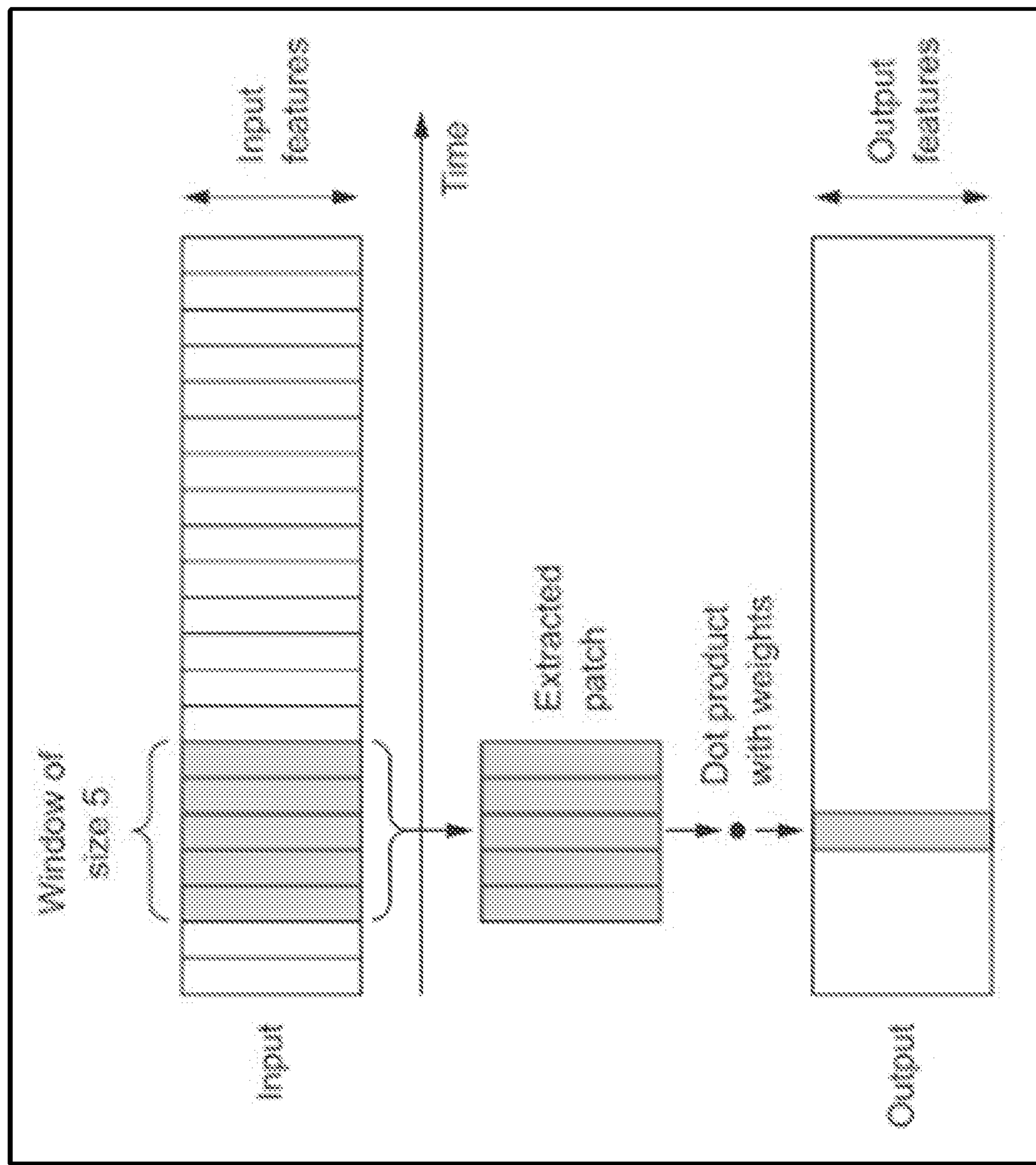


FIG. 1M

Global Average Pooling (GAP) in Convolutional Neural Networks

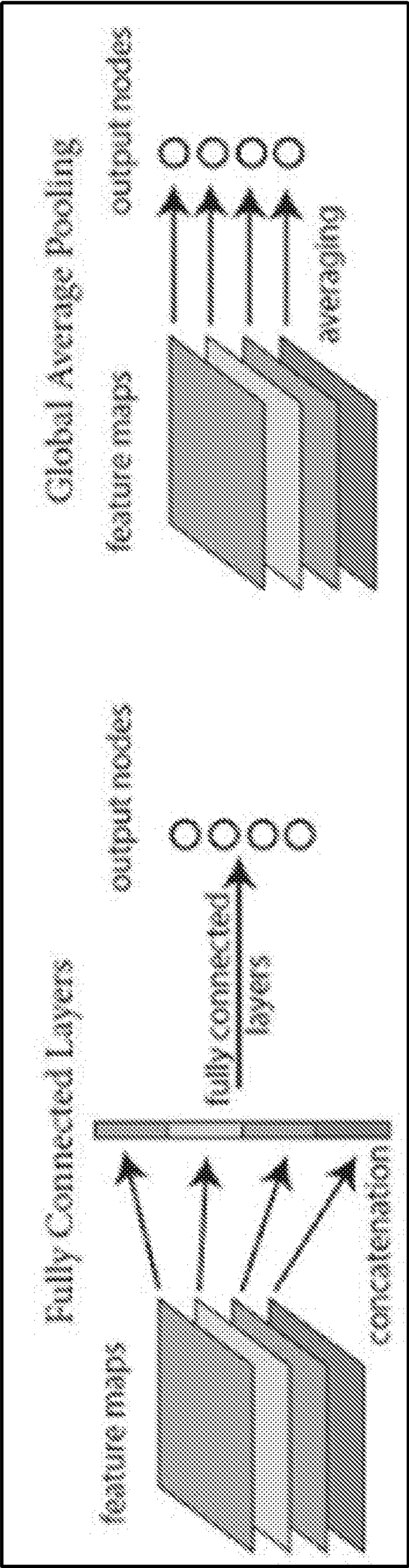
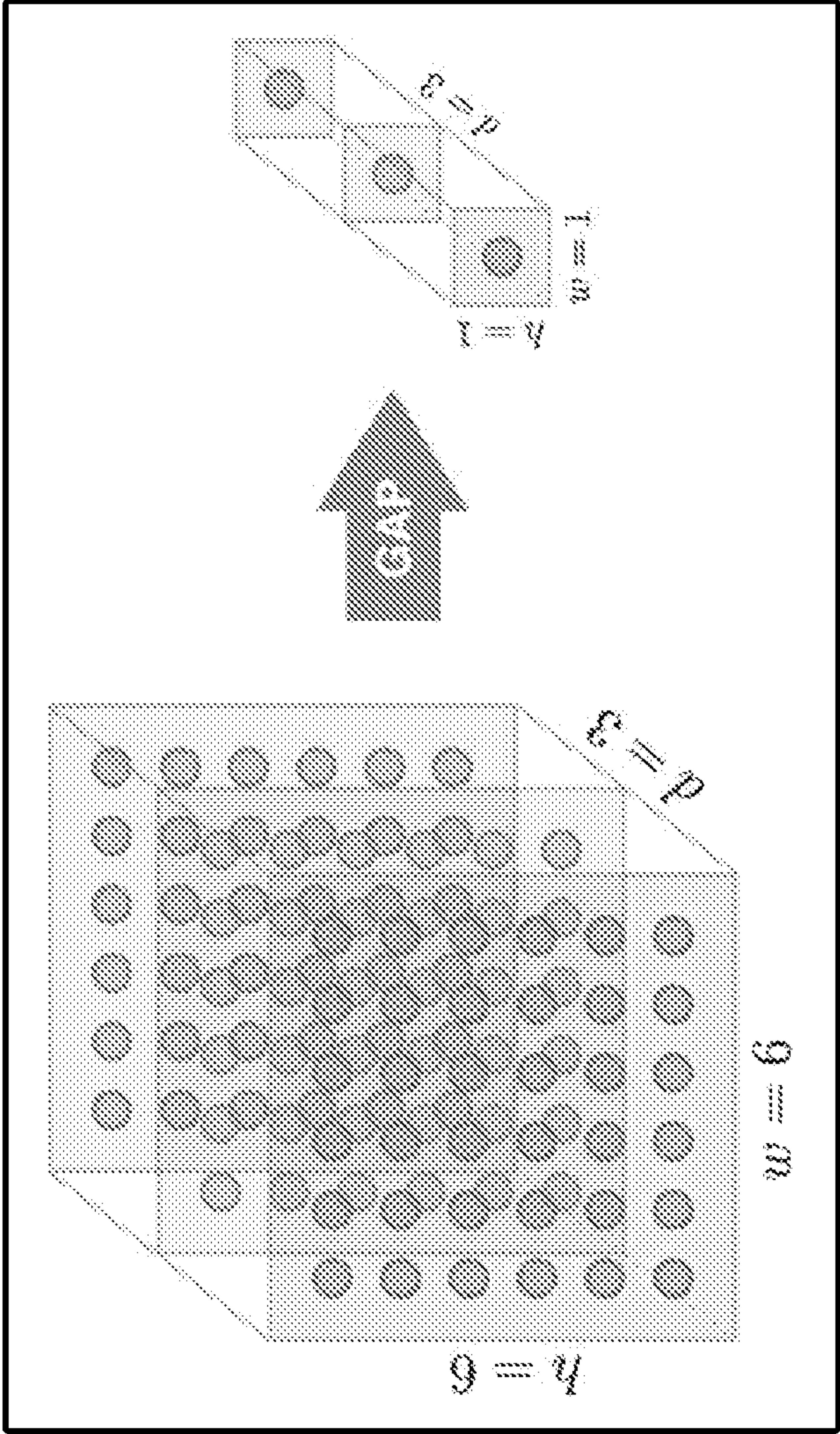


FIG. 1N

Dilated Convolutions

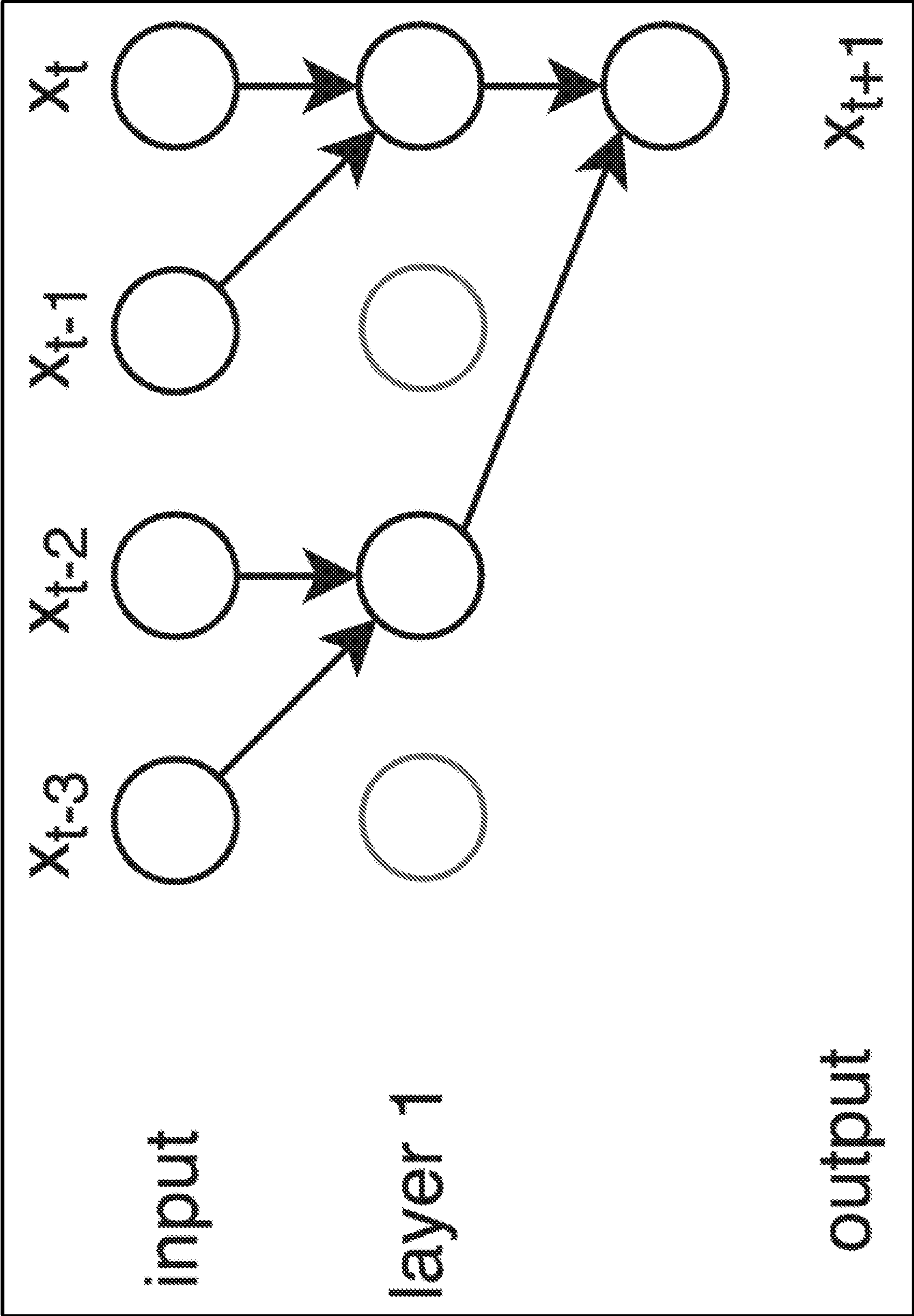


FIG. 10

Stacked Dilated Convolutions

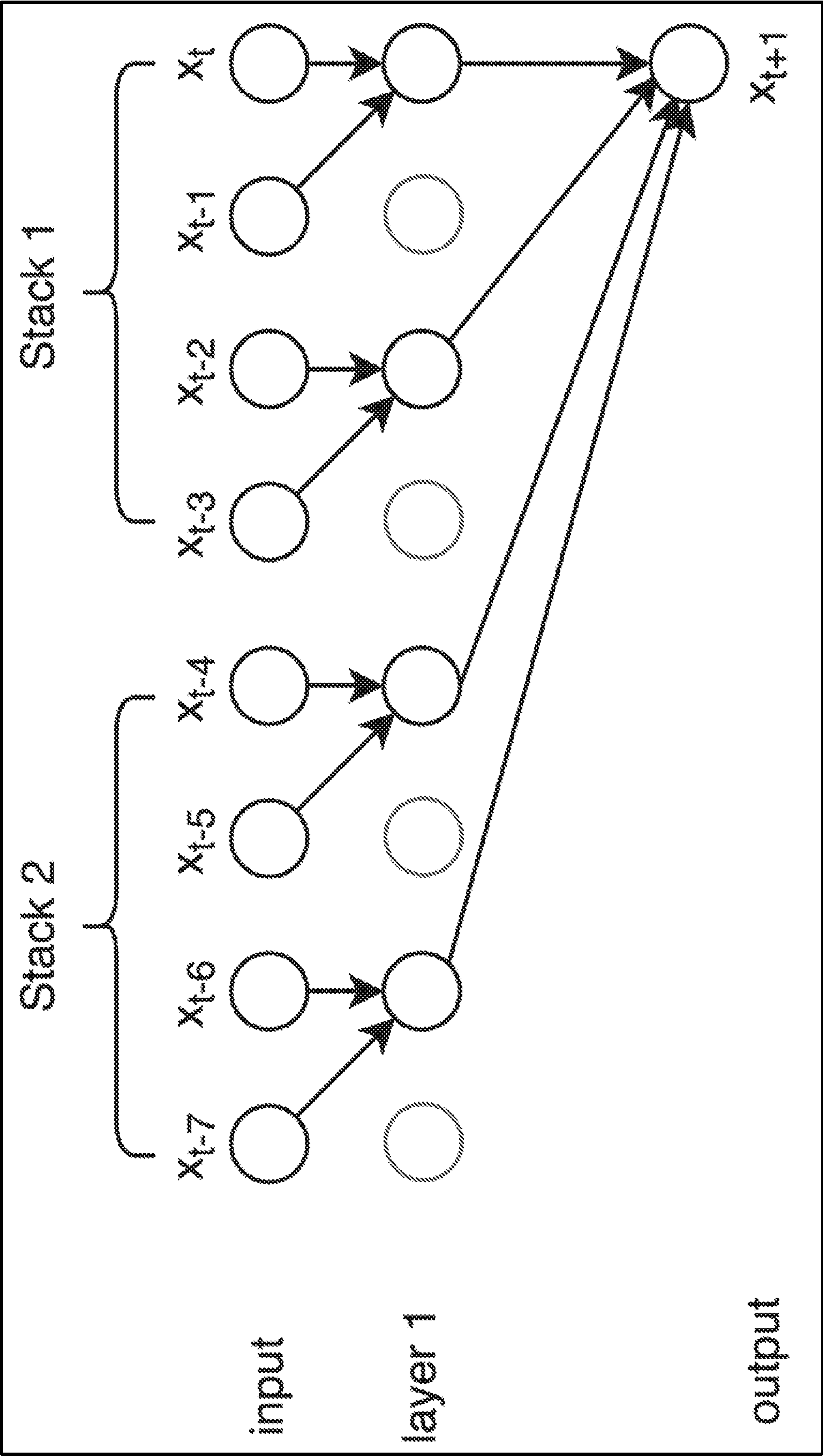


FIG. 1P

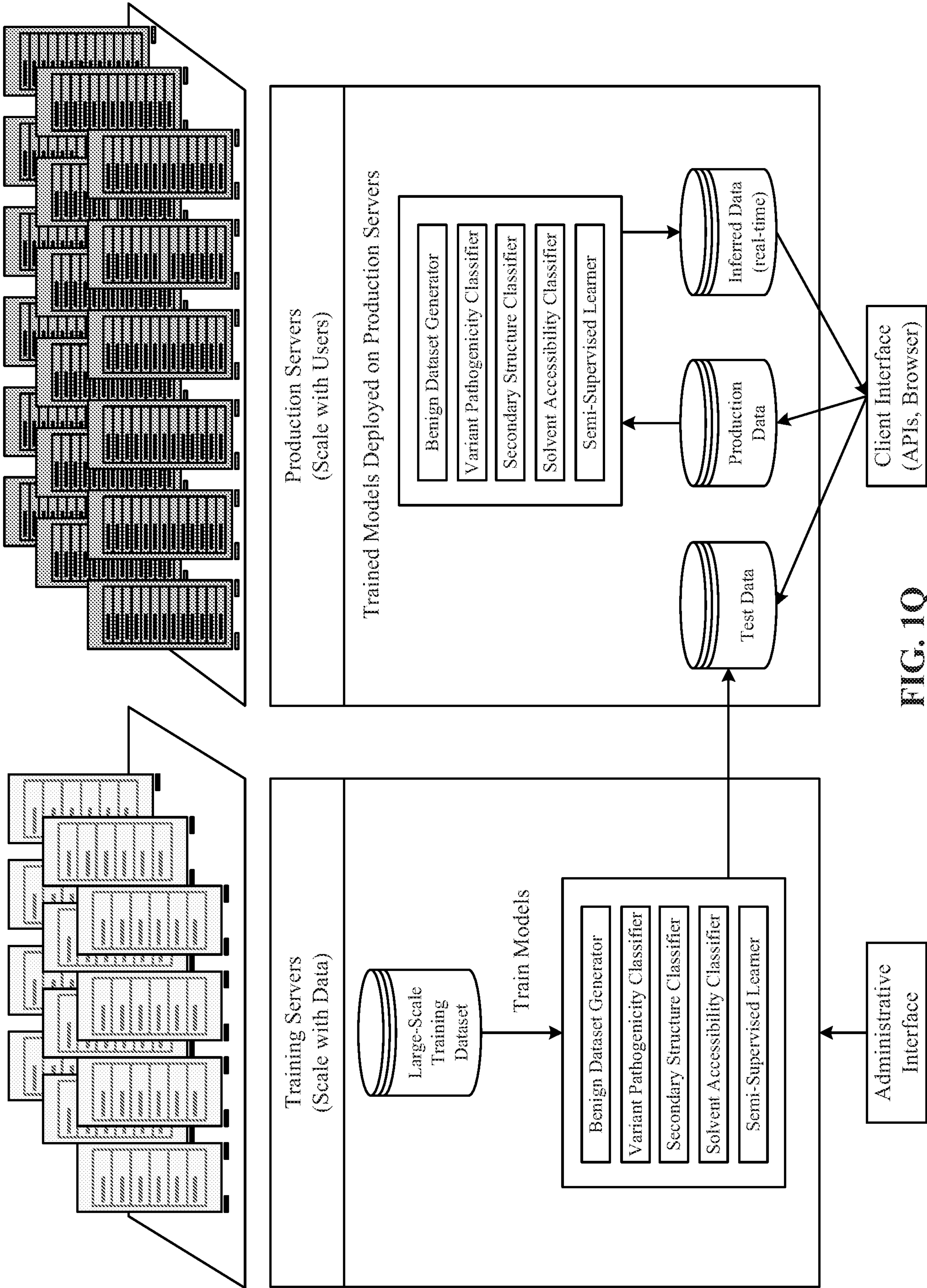


FIG. 1Q

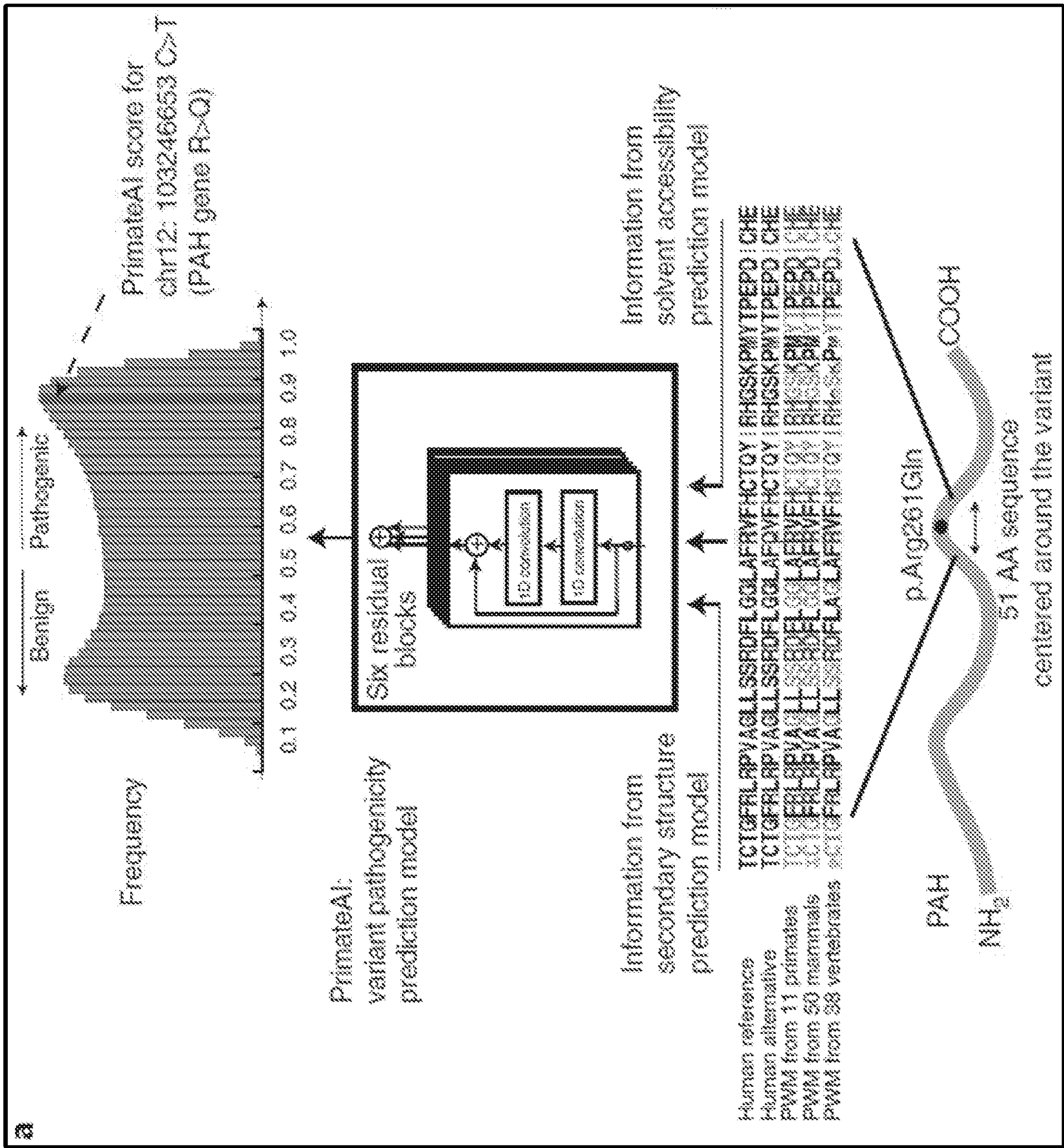
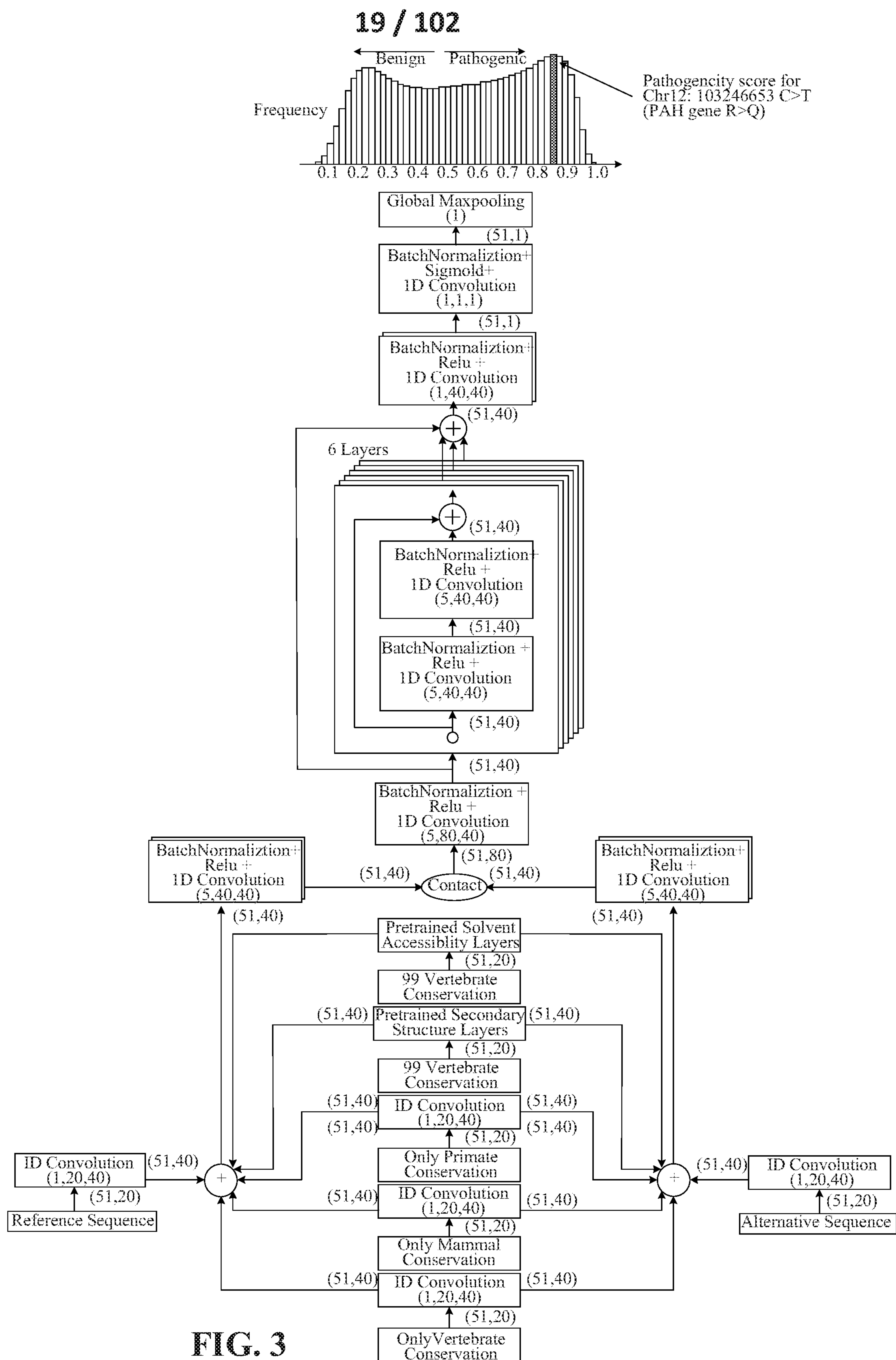


FIG. 2



Layer	Input to the Layer	Type	Number of Kernel & Window size	Shape	Activation
Layer 1a	Reference Sequence	Convolution 1D	40,1	(51,40)	Linear
Layer 1b	Alternative Sequence	Convolution 1D	40,1	(51,40)	Linear
Layer 1c	Primate Conservation	Convolution 1D	40,1	(51,40)	Linear
Layer 1d	Mammal Conservation	Convolution 1D	40,1	(51,40)	Linear
Layer 1e	Vertebrate Conservation	Convolution 1D	40,1	(51,40)	Linear
Pre-trained Secondary Structure Prediction Model (Trainable Network)	Primate + Mammal + Vertebrate Conservation	-	-	(51,40)	-
Pre-trained Subunit Accumulation Model (Trainable Network)	Primate + Mammal + Vertebrate Conservation	-	-	(51,40)	-
Layer 1f	Pre-trained Secondary Structure Prediction Model output	Convolution 1D	40,1	(51,40)	Linear
Layer 1g	Pre-trained Subunit Accumulation Model output	Convolution 1D	40,1	(51,40)	Linear
Merge 1a	Layer 1a, Layer 1c, Layer 1d, Layer 1e, Layer 1f, Layer 1g	Add	-	(51,40)	-
Merge 1b	Layer 1b, Layer 1c, Layer 1d, Layer 1e, Layer 1f, Layer 1g	Add	-	(51,40)	-
Layer 2a	Merge 1a	Batch Normalization	-	(51,40)	-
Layer 3a	Layer 2a	Activation	-	(51,40)	Relu
Layer 4a	Layer 3a	Convolution 1D	40,5	(51,40)	Linear
Layer 5a	Layer 4a	Batch Normalization	-	(51,40)	-
Layer 6a	Layer 5a	Activation	-	(51,40)	Relu
Layer 7a	Layer 6a	Convolution 1D	40,5	(51,40)	Linear
Layer 2b	Merge 1b	Batch Normalization	-	(51,40)	-
Layer 3b	Layer 2b	Activation	-	(51,40)	Relu
Layer 4b	Layer 3b	Convolution 1D	40,5	(51,40)	Linear

FIG. 4A

Layer 4b	Layer 4b	Batch Normalization	-	(\$1.40)	-
Layer 5b	Layer 5b	Activation	-	(\$1.40)	Relu
Layer 7b	Layer 6b	Convolution 1D	40.3	(\$1.40)	Linear
Merge 2a	Layer 7a, Layer 7b	Concatenate	-	(\$1.80)	-
Merge 2b	Layer 7a, Layer 7b	Concatenate	-	(\$1.80)	-
Layer 8a	Merge 2a	Convolution 1D	40.3	(\$1.40)	Linear
Layer 8b	Merge 2b	Convolution 1D	40.3	(\$1.40)	Linear
Layer 9	Layer 8a	Batch Normalization	-	(\$1.40)	-
Layer 10	Layer 9	Activation	-	(\$1.40)	Relu
Layer 11	Layer 10	Convolution 1D	40.3	(\$1.40)	Linear
Layer 12	Layer 11	Batch Normalization	-	(\$1.40)	-
Layer 13	Layer 12	Activation	-	(\$1.40)	Relu
Layer 14	Layer 13	Convolution 1D	40.3	(\$1.40)	Linear
Merge3	Layer 8a, Layer 14	Add	-	(\$1.40)	-
Layer 15	Merge3	Batch Normalization	-	(\$1.40)	-
Layer 16	Layer 15	Activation	-	(\$1.40)	Relu
Layer 17	Layer 16	Convolution 1D	40.3	(\$1.40)	Linear
Layer 18	Layer 17	Batch Normalization	-	(\$1.40)	-
Layer 19	Layer 18	Activation	-	(\$1.40)	Relu
Layer 20	Layer 19	Convolution 1D	40.3	(\$1.40)	Linear
Merge4	Merge3, Layer 20	Add	-	(\$1.40)	-
Layer 21	Merge4	Convolution 1D	40.3	(\$1.40)	Linear
Layer 22	Merge4	Batch Normalization	-	(\$1.40)	-
Layer 23	Layer 22	Activation	-	(\$1.40)	Relu
Layer 24	Layer 23	Convolution 1D	40.3	(\$1.40)	Linear
Layer 25	Layer 24	Batch Normalization	-	(\$1.40)	-
Layer 26	Layer 25	Activation	-	(\$1.40)	Relu
Layer 27	Layer 26	Convolution 1D	40.3	(\$1.40)	Linear
Merge5	Merge4, Layer 27	Add	-	(\$1.40)	-
Layer 28	Merge 5	Batch Normalization	-	(\$1.40)	-
Layer 29	Layer 28	Activation	-	(\$1.40)	Relu
Layer 30	Layer 29	Convolution 1D	40.3	(\$1.40)	Linear
Layer 31	Layer 30	Batch Normalization	-	(\$1.40)	-
Layer 32	Layer 31	Activation	-	(\$1.40)	Relu

FIG. 4B

Layer 33	Layer 32	Convolution 1D	40,5	(51,40)	Linear
Merge 6	Merge 5, Layer 33	Add	-	(51,40)	-
Layer 34	Merge 6	Convolution 1D	40,1	(51,40)	Linear
Layer 35	Merge 6	Batch Normalization	-	(51,40)	-
Layer 36	Layer 35	Activation	-	(51,40)	Relu
Layer 37	Layer 36	Convolution 1D	40,3	(51,40)	Linear
Layer 38	Layer 37	Batch Normalization	-	(51,40)	-
Layer 39	Layer 38	Activation	-	(51,40)	Relu
Layer 40	Layer 39	Convolution 1D	40,5	(51,40)	Linear
Merge 7	Merge 6, Layer 40	Add	-	(51,40)	-
Layer 41	Merge 7	Batch Normalization	-	(51,40)	-
Layer 42	Layer 41	Activation	-	(51,40)	Relu
Layer 43	Layer 42	Convolution 1D	40,5	(51,40)	Linear
Layer 44	Layer 43	Batch Normalization	-	(51,40)	-
Layer 45	Layer 44	Activation	-	(51,40)	Relu
Layer 46	Layer 45	Convolution 1D	40,3	(51,40)	Linear
Merge 8	Merge 7, Layer 46	Add	-	(51,40)	-
Layer 47	Merge 8	Convolution 1D	40,3	(51,40)	Linear
Merge 9	Layer 38, Layer 21, Layer 34, Layer 47	Add	-	(51,40)	-
Layer 48	Merge 10	Batch Normalization	-	(51,40)	-
Layer 49	Layer 48	Activation	-	(51,40)	Relu
Layer 50	Layer 49	Convolution 1D	40,3	(51,40)	Linear
Layer 51	Layer 50	Batch Normalization	-	(51,40)	-
Layer 52	Layer 51	Activation	-	(51,40)	Relu
Layer 53	Layer 52	Convolution 1D	40,1	(51,40)	Linear
Merge 10	Merge 9, Layer 53	Add	-	(51,40)	-
Layer 54	Merge 10	Convolution 1D	1,1	(51,1)	Sigmoid
Output layer	Layer 54	GlobalMaxPooling1D	-	-	-

FIG. 4C

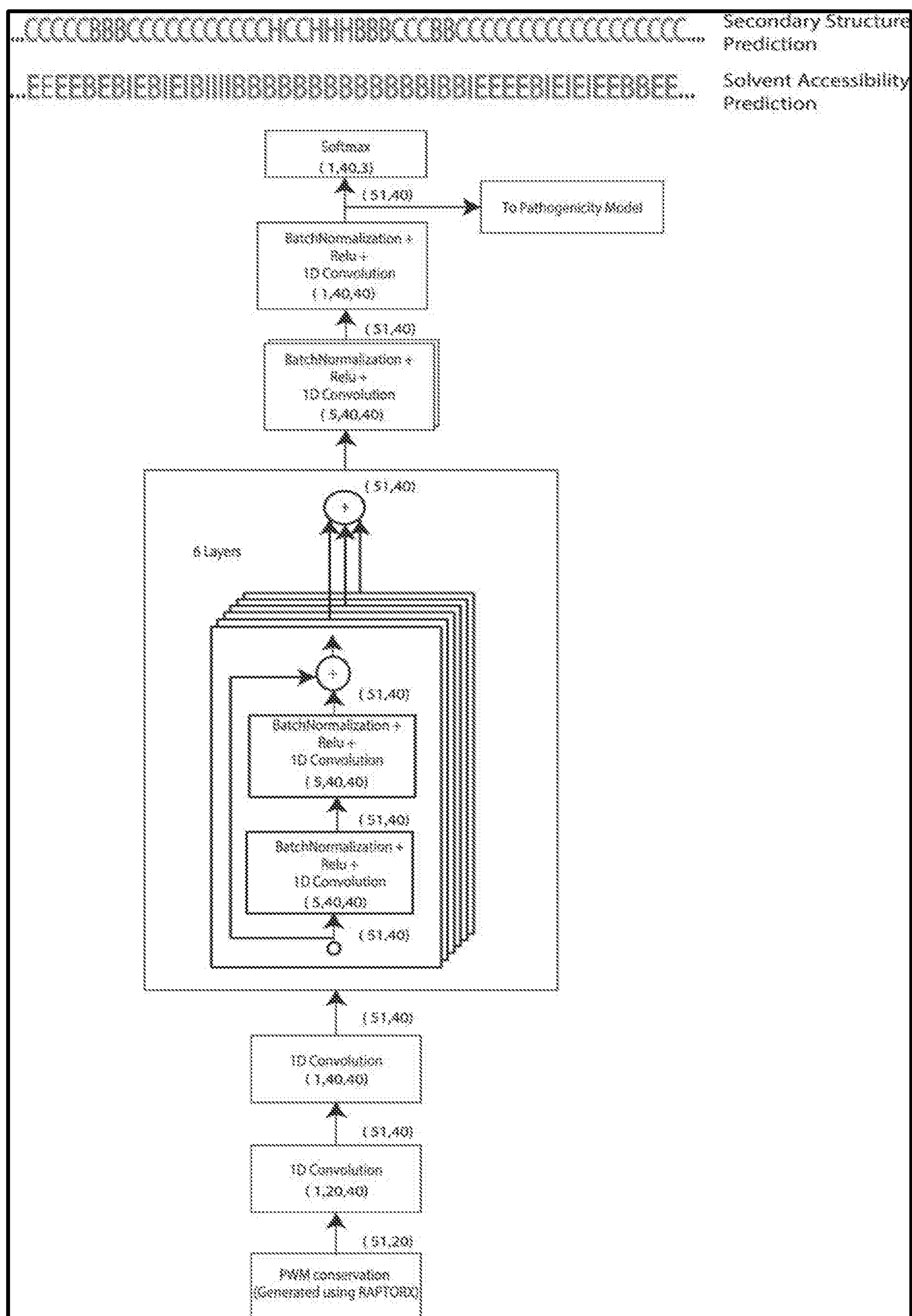


FIG. 6

Layer	Input to the Layer	Type	Number of Kernel & Kernel Window size	Shape	Activation
Layer 1a	Input PSFM	Convolution 1D	40,1	(51,40)	Linear
Layer 2a	Layer 1a	Convolution 1D	40,1	(51,40)	Linear
Layer 1b	Input PSFM	Convolution 1D	40,1	(51,40)	Linear
Layer 2b	Layer 1b	Convolution 1D	40,1	(51,40)	Linear
Layer 3	Layer 2a	Batch Normalization	-	(51,40)	-
Layer 4	Layer 3	Activation	-	(51,40)	Relu
Layer 5	Layer 4	Convolution 1D	40,5	(51,40)	Linear
Layer 6	Layer 5	Batch Normalization	-	(51,40)	-
Layer 8	Layer 7	Activation	-	(51,40)	Relu
Layer 9	Layer 8	Convolution 1D	40,5	(51,40)	Linear
Merge1	Layer 9, Layer 9	Add	-	(51,40)	-
Layer 10	Merge1	Batch Normalization	-	(51,40)	-
Layer 11	Layer 10	Activation	-	(51,40)	Relu
Layer 12	Layer 11	Convolution 1D	40,5	(51,40)	Linear
Layer 13	Layer 12	Batch Normalization	-	(51,40)	-
Layer 14	Layer 13	Activation	-	(51,40)	Relu
Layer 15	Layer 14	Convolution 1D	40,5	(51,40)	Linear
Merge2	Merge1, Layer15	Add	-	(51,40)	-
Layer 16	Merge2	Convolution 1D	40,1	(51,40)	Linear
Layer 17	Merge2	Batch Normalization	-	(51,40)	-
Layer 18	Layer 17	Activation	-	(51,40)	Relu
Layer 19	Layer 18	Convolution 1D	40,5	(51,40)	Linear
Layer 20	Layer 19	Batch Normalization	-	(51,40)	-
Layer 21	Layer 20	Activation	-	(51,40)	Relu
Layer 22	Layer 21	Convolution 1D	40,5	(51,40)	Linear
Merge3	Merge1, Layer 22	Add	-	(51,40)	-

FIG. 7A

Layer 23	Merge 3	Batch Normalization	-	(51,40)	-
Layer 24	Layer 23	Activation	-	(51,40)	Relu
Layer 25	Layer 24	Convolution 1D	40,5	(51,40)	Linear
Layer 26	Layer 25	Batch Normalization	-	(51,40)	-
Layer 27	Layer 26	Activation	-	(51,40)	Relu
Layer 28	Layer 27	Convolution 1D	40,5	(51,40)	Linear
Merge4	Merge3, Layer 28	Add	-	(51,40)	-
Layer 29	Merge4	Convolution 1D	40,1	(51,40)	Linear
Layer 30	Merge4	Batch Normalization	-	(51,40)	-
Layer 31	Layer 30	Activation	-	(51,40)	Relu
Layer 32	Layer 31	Convolution 1D	40,5	(51,40)	Linear
Layer 33	Layer 32	Batch Normalization	-	(51,40)	-
Layer 34	Layer 33	Activation	-	(51,40)	Relu
Layer 35	Layer 34	Convolution 1D	40,5	(51,40)	Linear
Merge5	Merge4, Layer 35	Add	-	(51,40)	-
Layer 36	Merge5	Batch Normalization	-	(51,40)	-
Layer 37	Layer 36	Activation	-	(51,40)	Relu
Layer 38	Layer 37	Convolution 1D	40,5	(51,40)	Linear
Layer 39	Layer 38	Batch Normalization	-	(51,40)	-
Layer 40	Layer 39	Activation	-	(51,40)	Relu
Layer 41	Layer 40	Convolution 1D	40,5	(51,40)	Linear
Merge6	Merge5, Layer 41	Add	-	(51,40)	-
Layer 42	Merge 6	Convolution 1D	40,1	(51,40)	Linear
Merge7	Layer 23, Layer 15 Layer 39, Layer 42	Add	-	(51,40)	-
Layer 43	Merge 7	Batch Normalization	-	(51,40)	-
Layer 44	Layer 43	Activation	-	(51,40)	Relu
Layer 45	Merge 44	Convolution 1D	40,1	(51,40)	Linear
Layer 46	Layer 45	Batch Normalization	-	(51,40)	-
Layer 47	Layer 46	Activation	-	(51,40)	Relu
Layer 48	Layer 47	Convolution 1D	40,1	(51,40)	Linear
Merge 8	Merge 7, Layer 48	Add	-	(51,40)	-
Output layer	Merge 8	Convolution 1D	3,1	(51,5)	Softmax

FIG. 7B

Layer	Input to the Layer	Type	Number of Kernel & Kernel Window size	Shape	Activation
Layer 1a	Input PAFM	Convolution 1D	40,1	(51,40)	Linear
Layer 2a	Layer 1a	Convolution 1D	40,1	(51,40)	Linear
Layer 1b	Input PAFM	Convolution 1D	40,1	(51,40)	Linear
Layer 2b	Layer 1b	Convolution 1D	40,1	(51,40)	Linear
Layer3	Layer 2a	Batch Normalization	-	(51,40)	-
Layer 4	Layer 3	Activation	-	(51,40)	Relu
Layer 5	Layer 4	Convolution 1D	40,1	(51,40)	Linear
Layer6	Layer5	Batch Normalization	-	(51,40)	-
Layer8	Layer 7	Activation	-	(51,40)	Relu
Layer 9	Layer 8	Convolution 1D	40,5	(51,40)	Linear
Merge1	Layer 2a, Layer 9	Add	-	(51,40)	-
Layer 10	Merge1	Batch Normalization	-	(51,40)	-
Layer 11	Layer 10	Activation	-	(51,40)	Relu
Layer 12	Layer 11	Convolution 1D	40,5	(51,40)	Linear
Layer 13	Layer 12	Batch Normalization	-	(51,40)	-
Layer 14	Layer 13	Activation	-	(51,40)	Relu
Layer 15	Layer 14	Convolution 1D	40,5	(51,40)	Linear
Merge2	Merge1, Layer15	Add	-	(51,40)	-
Layer 16	Merge2	Convolution 1D	40,1	(51,40)	Linear
Layer 17	Merge2	Batch Normalization	-	(51,40)	-
Layer 18	Layer 17	Activation	-	(51,40)	Relu
Layer 19	Layer 18	Convolution 1D	40,5	(51,40)	Linear
Layer 20	Layer 19	Batch Normalization	-	(51,40)	-
Layer 21	Layer 20	Activation	-	(51,40)	Relu
Layer 22	Layer 21	Convolution 1D	40,5	(51,40)	Linear
Merge3	Merge2, Layer 22	Add	-	(51,40)	-

FIG. 8A

Layer 23	Merge 3	Batch Normalization	-	(51,40)	-
Layer 24	Layer 23	Activation	-	(51,40)	Relu
Layer 25	Layer 24	Convolution 1D	40,5	(51,40)	Linear
Layer 26	Layer 25	Batch Normalization	-	(51,40)	-
Layer 27	Layer 26	Activation	-	(51,40)	Relu
Layer 28	Layer 27	Convolution 1D	40,5	(51,40)	Linear
Merge4	Merge3, Layer 28	Add	-	(51,40)	-
Layer 29	Merge4	Convolution 1D	40,1	(51,40)	Linear
Layer 30	Merge4	Batch Normalization	-	(51,40)	-
Layer 31	Layer 30	Activation	-	(51,40)	Relu
Layer 32	Layer 31	Convolution 1D	40,5	(51,40)	Linear
Layer 33	Layer 32	Batch Normalization	-	(51,40)	-
Layer 34	Layer 33	Activation	-	(51,40)	Relu
Layer 35	Layer 34	Convolution 1D	40,5	(51,40)	Linear
Merge5	Merge4, Layer 35	Add	-	(51,40)	-
Layer 36	Merge5	Batch Normalization	-	(51,40)	-
Layer 37	Layer 36	Activation	-	(51,40)	Relu
Layer 38	Layer 37	Convolution 1D	40,5	(51,40)	Linear
Layer 39	Layer 38	Batch Normalization	-	(51,40)	-
Layer 40	Layer 39	Activation	-	(51,40)	Relu
Layer 41	Layer 40	Convolution 1D	40,5	(51,40)	Linear
Merge6	Merge5, Layer 41	Add	-	(51,40)	-
Layer 42	Merge 6	Convolution 1D	40,1	(51,40)	Linear
Merge7	Layer 36, Layer 42 Layer 37, Layer 42	Add	-	(51,40)	-
Layer 43	Merge 7	Batch Normalization	-	(51,40)	-
Layer 44	Layer 43	Activation	-	(51,40)	Relu
Layer 45	Merge 44	Convolution 1D	40,1	(51,40)	Linear
Layer 46	Layer 45	Batch Normalization	-	(51,40)	-
Layer 47	Layer 46	Activation	-	(51,40)	Relu
Layer 48	Layer 47	Convolution 1D	40,1	(51,40)	Linear
Merge 8	Merge 7, Layer 48	Add	-	(51,40)	-
Output layer	Merge 8	Convolution 1D	3,1	(51,3)	Softmax

FIG. 8B

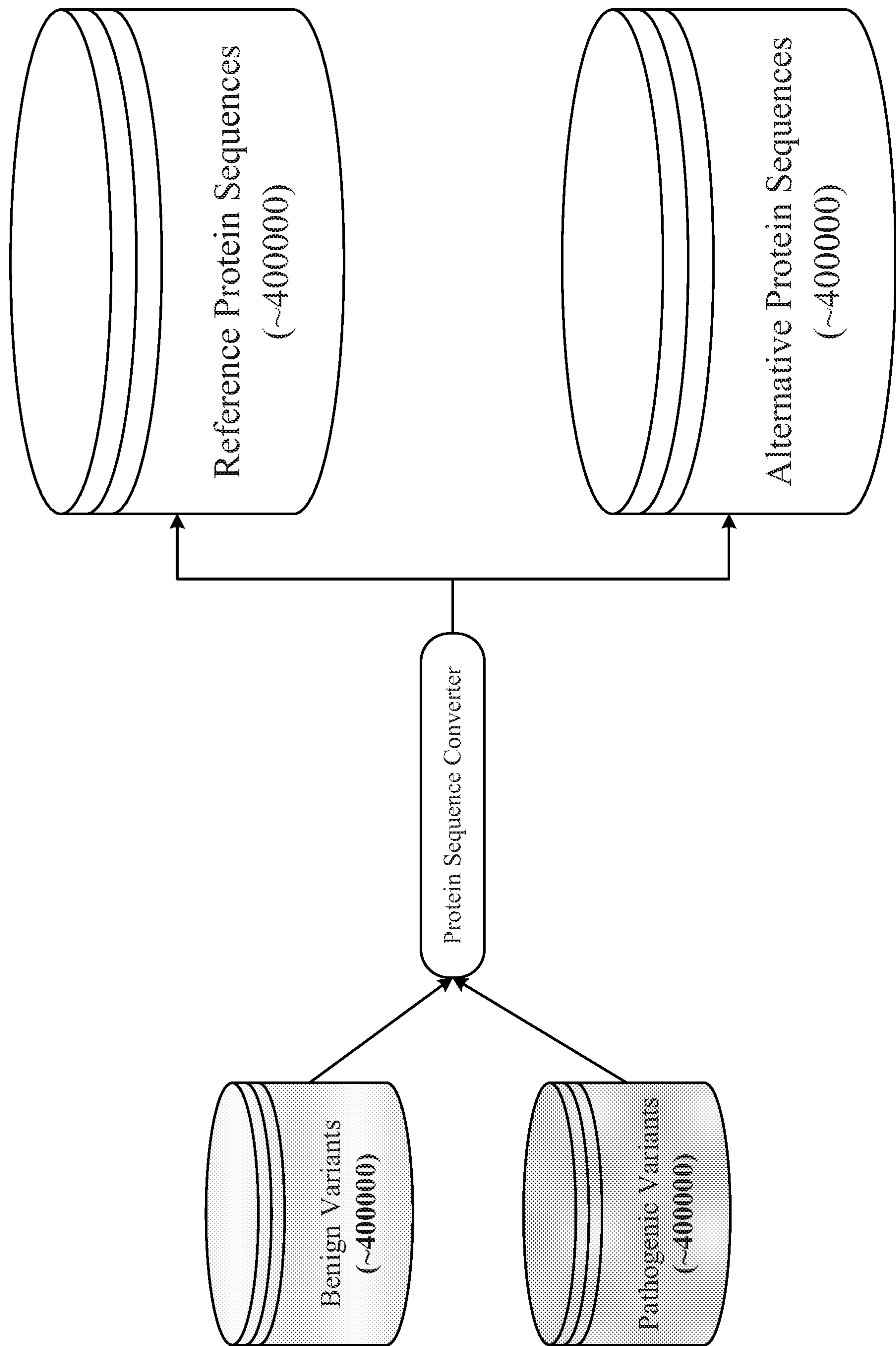


FIG. 9

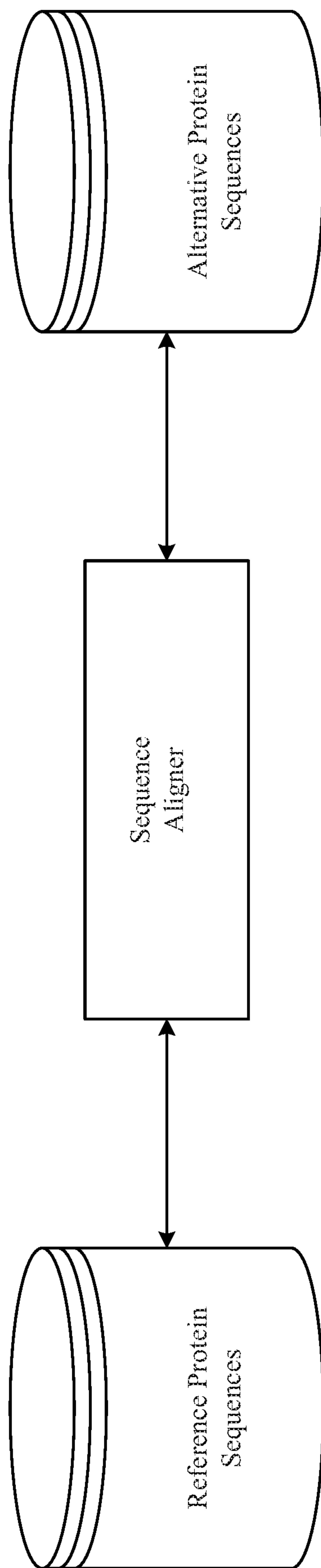


FIG. 10

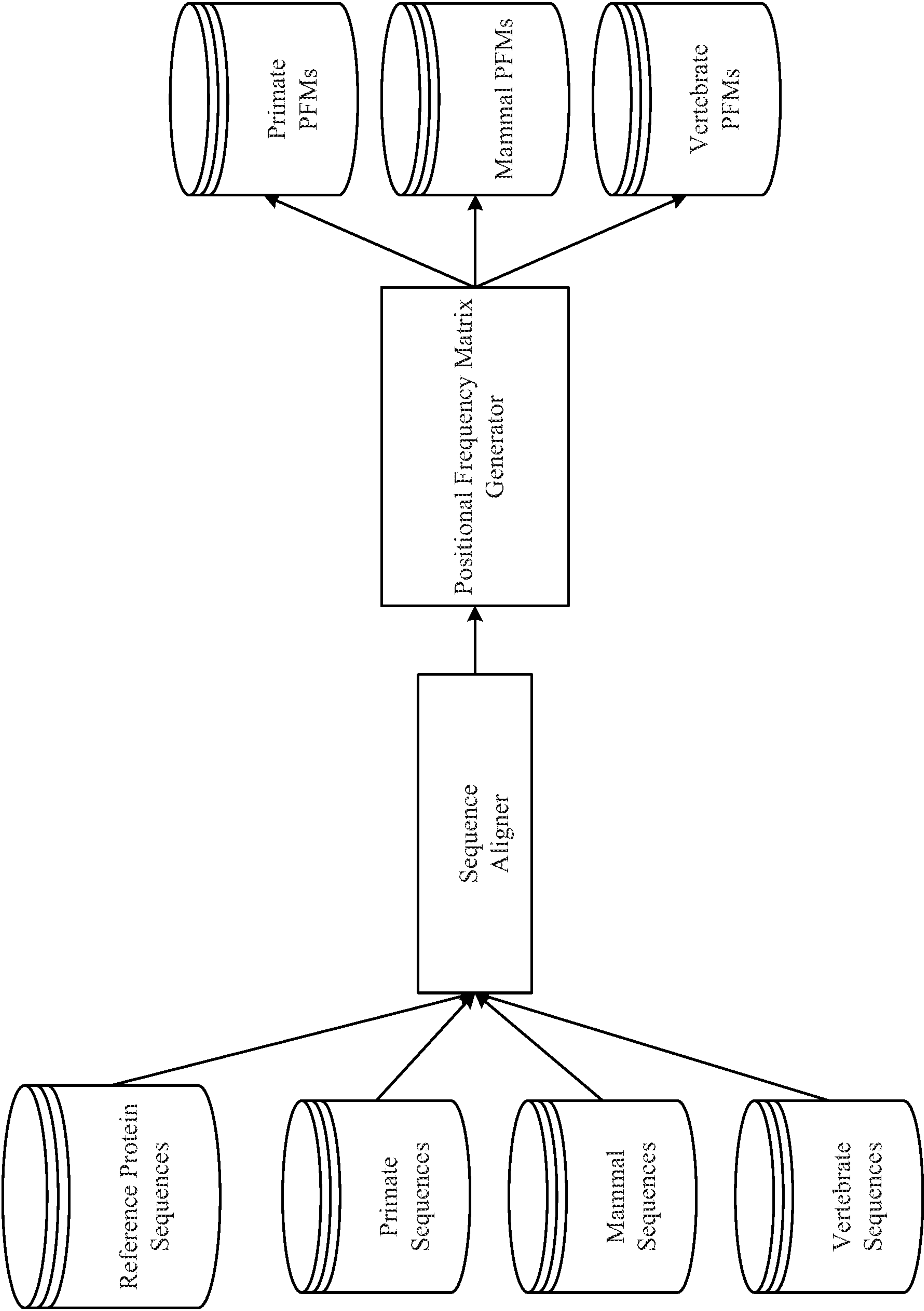


FIG. 11

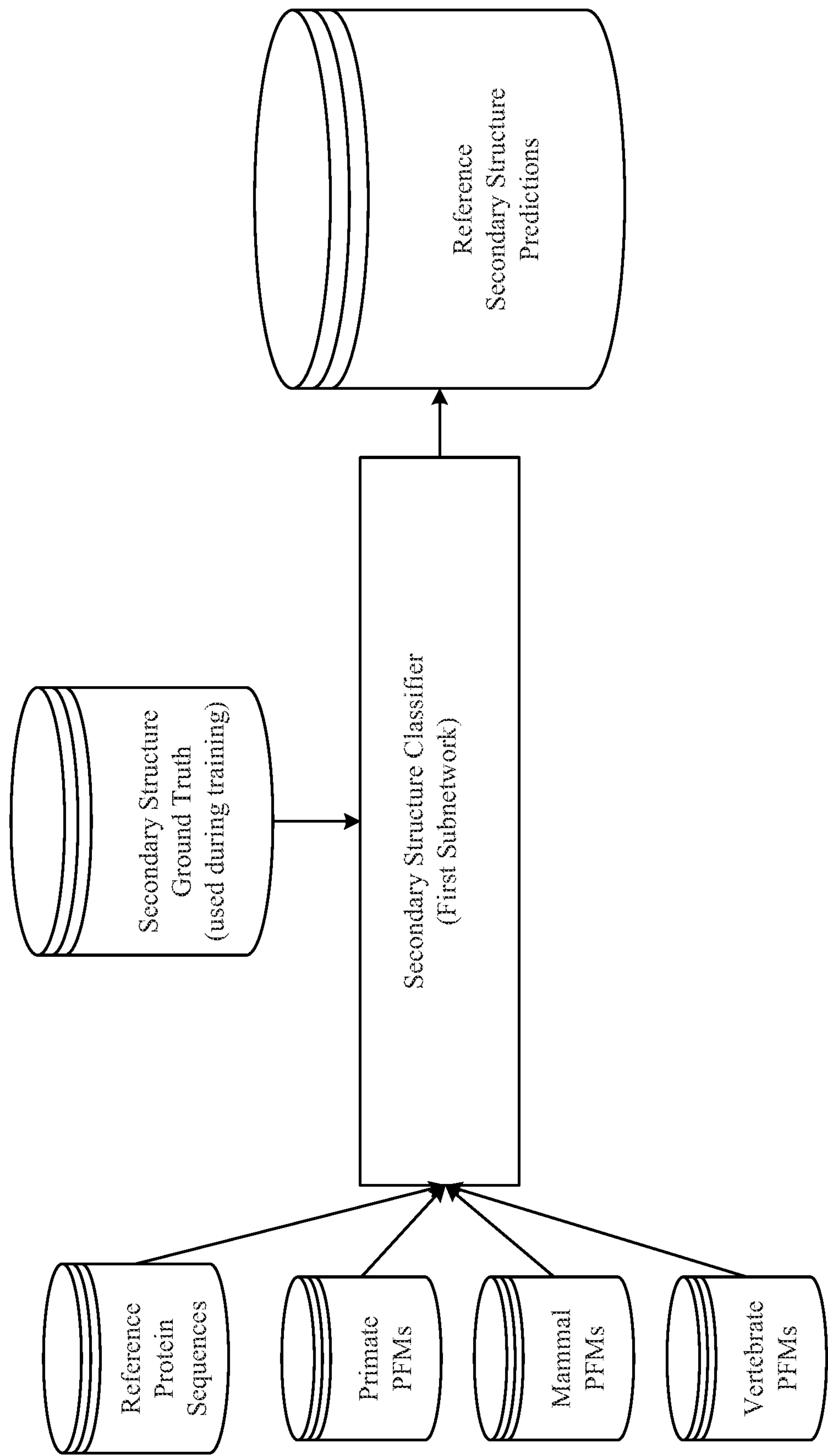


FIG. 12

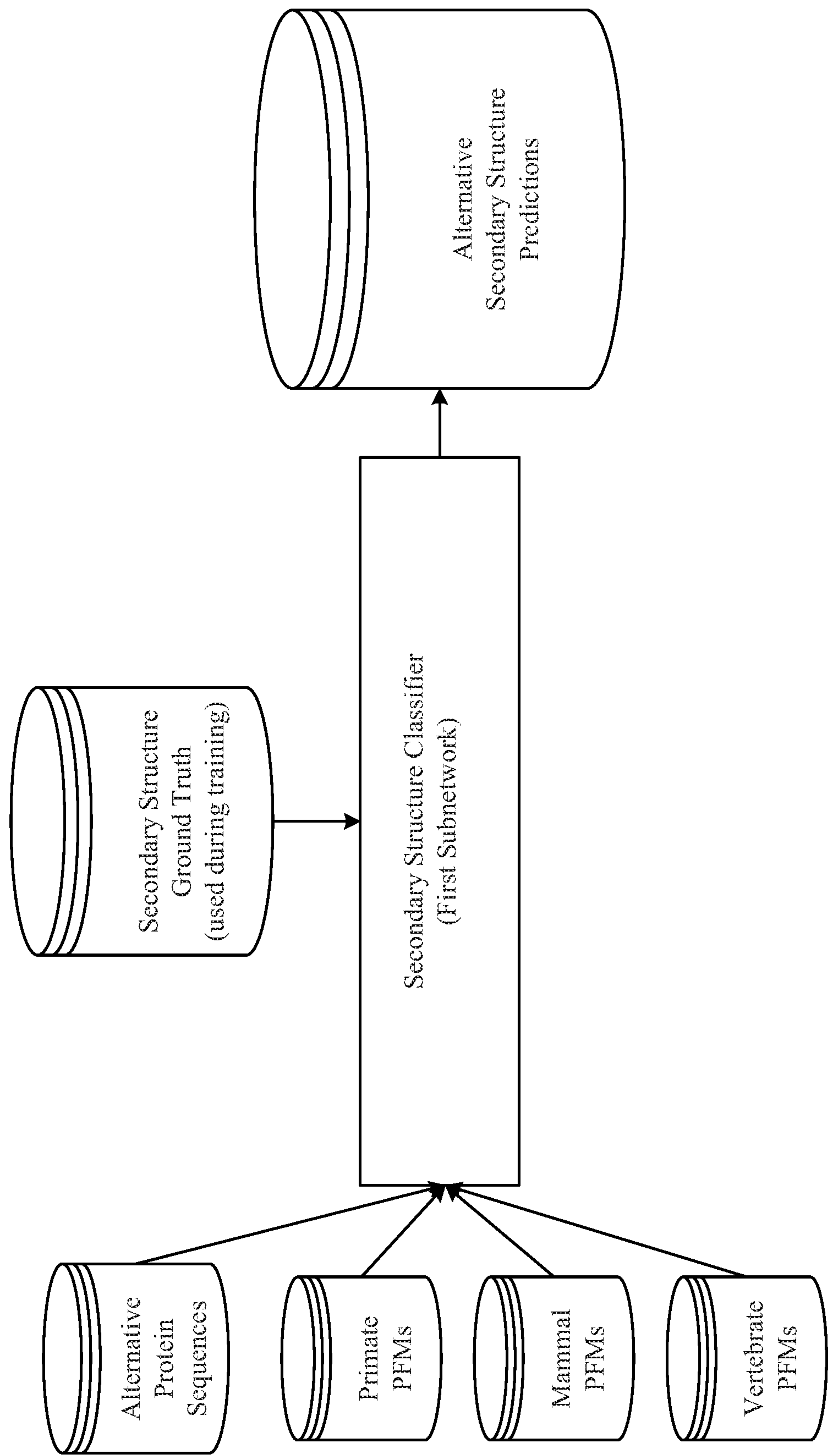


FIG. 13

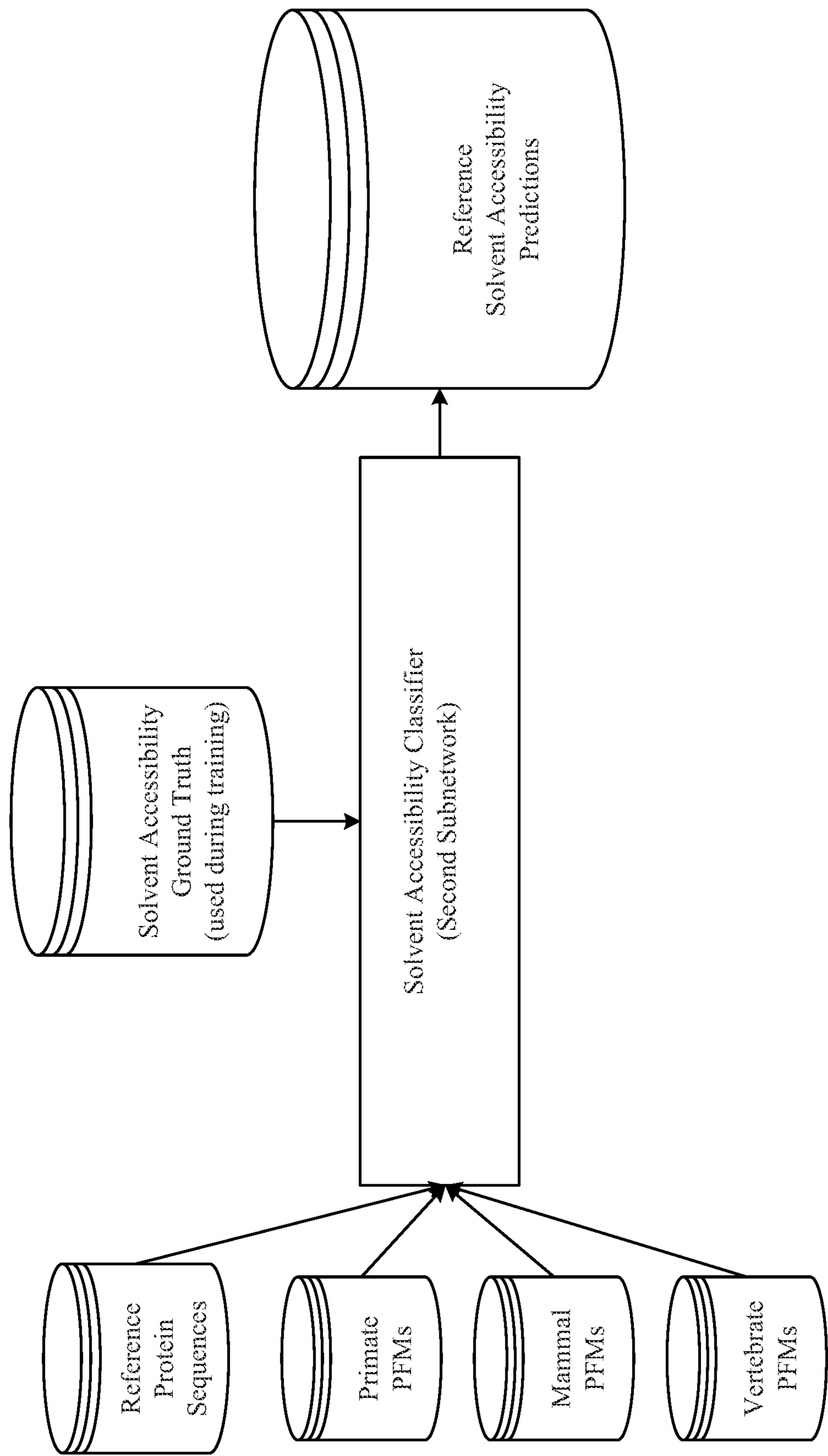


FIG. 14

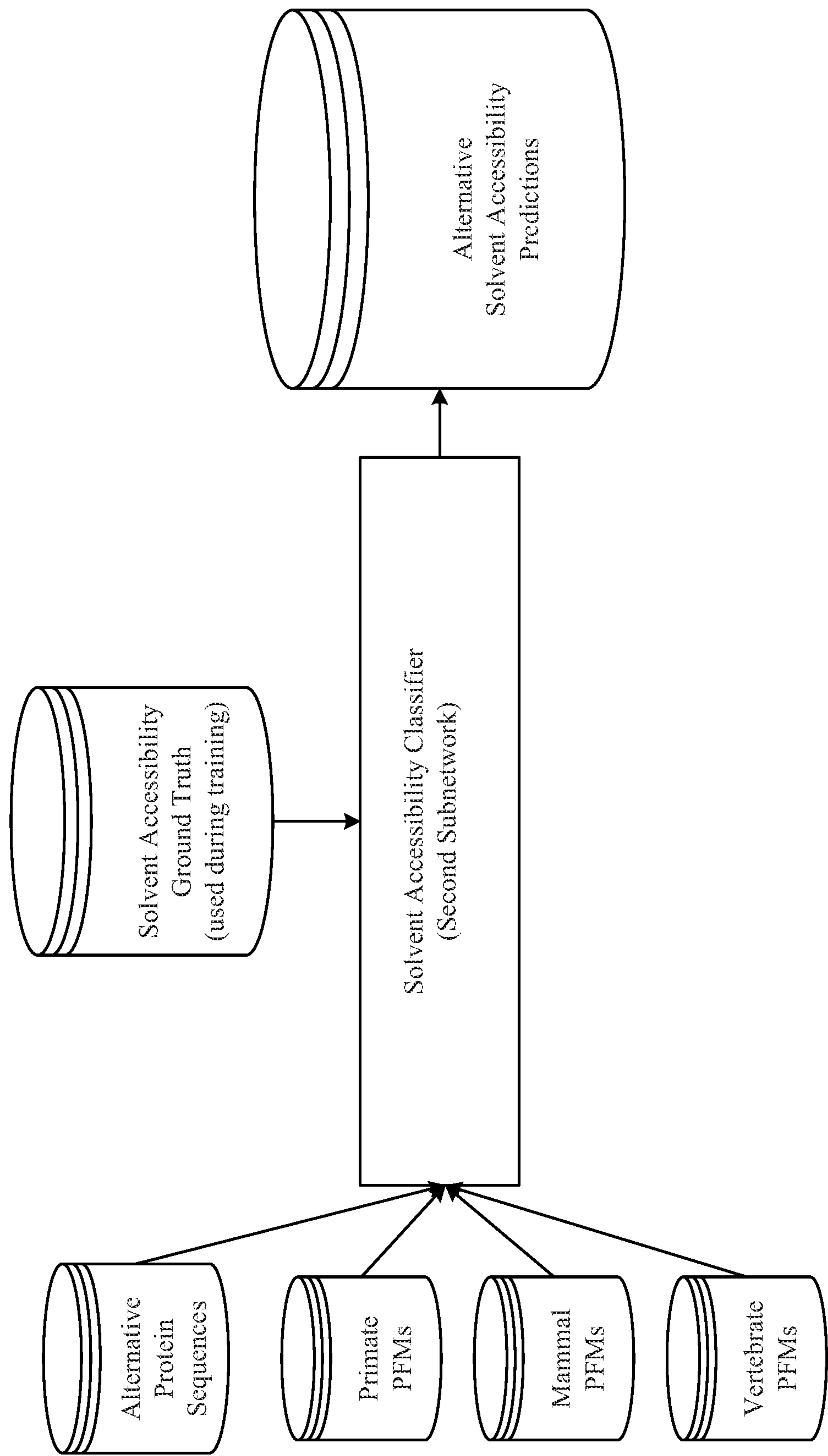


FIG. 15

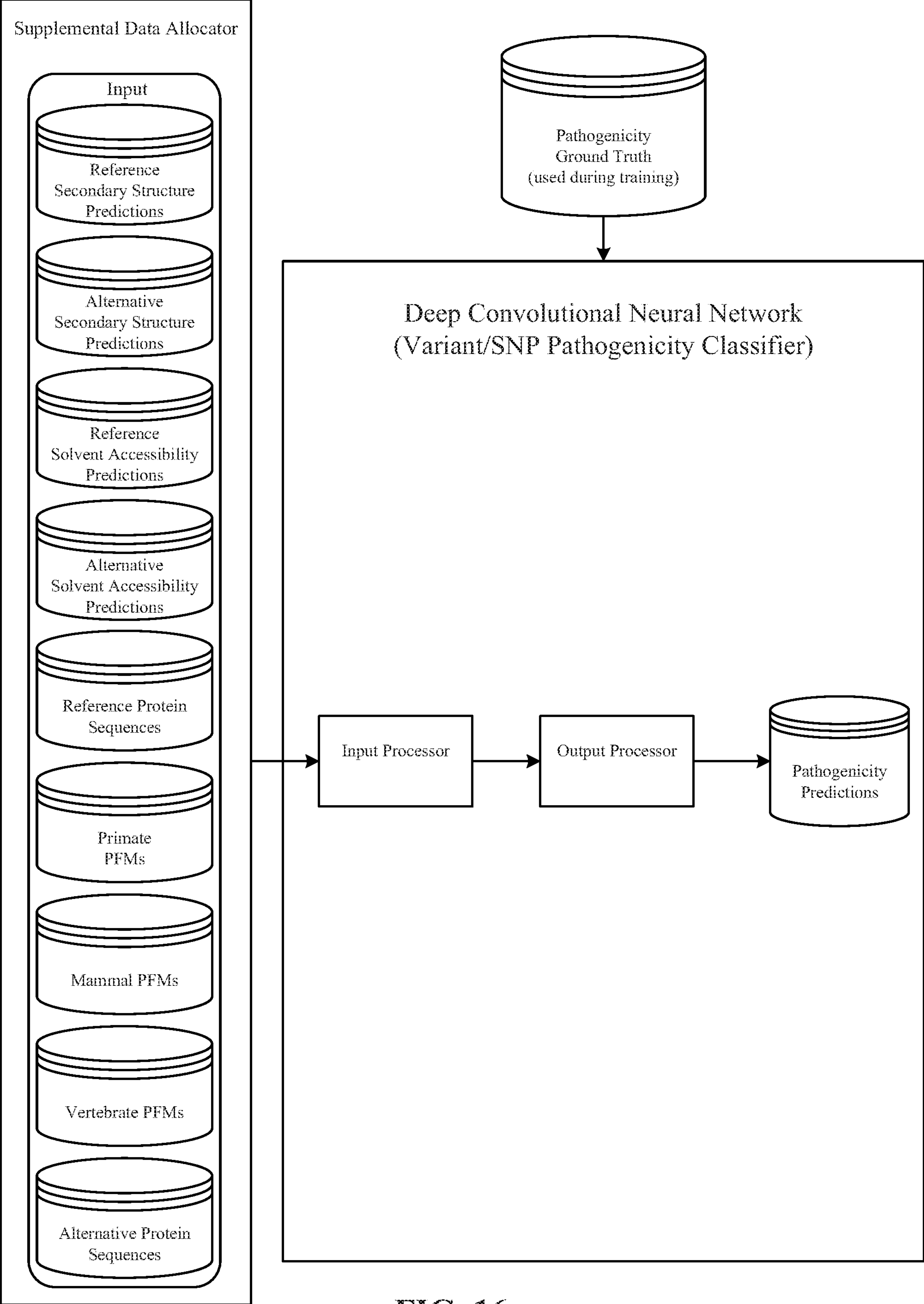
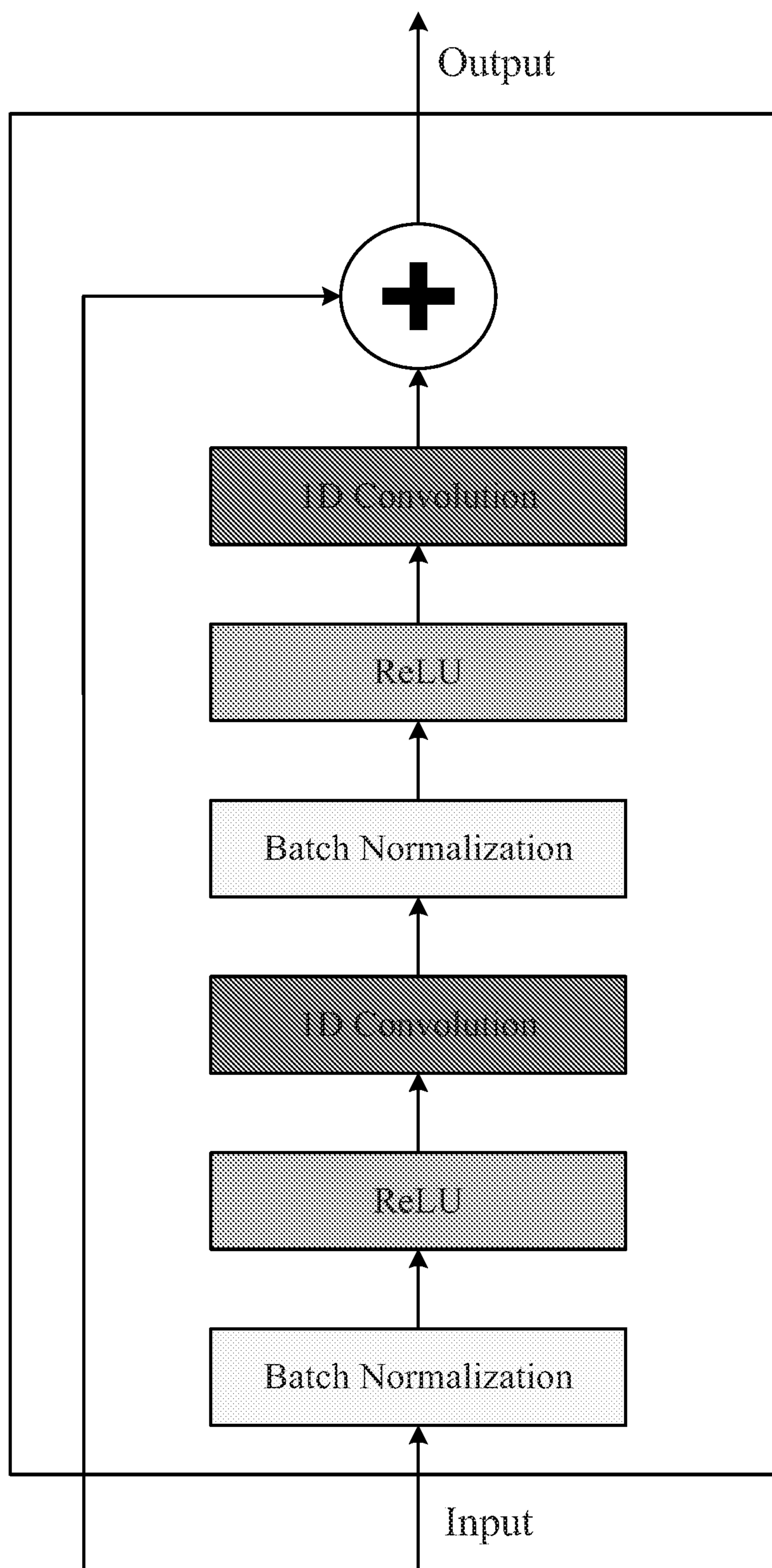


FIG. 16

**FIG. 17**

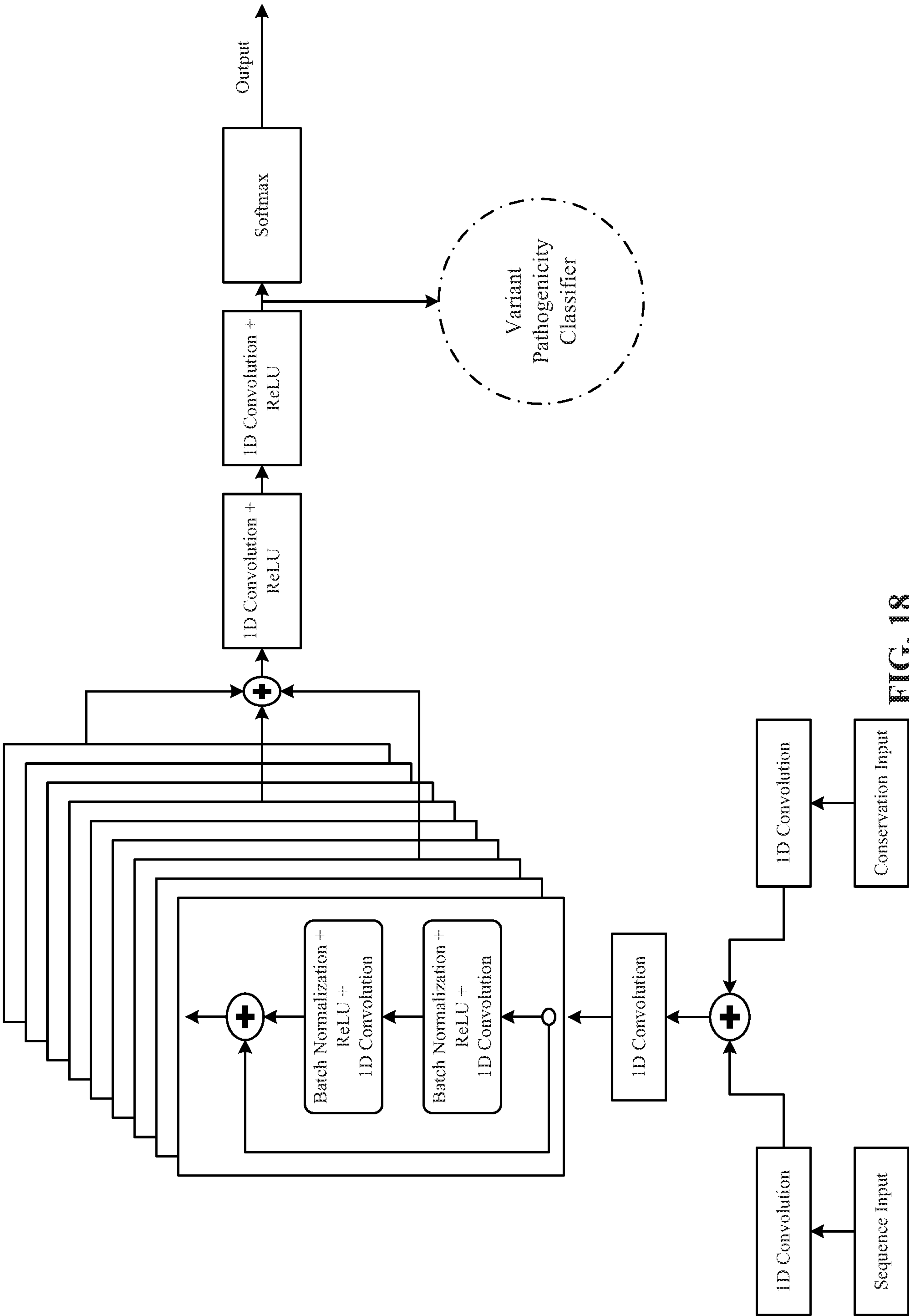
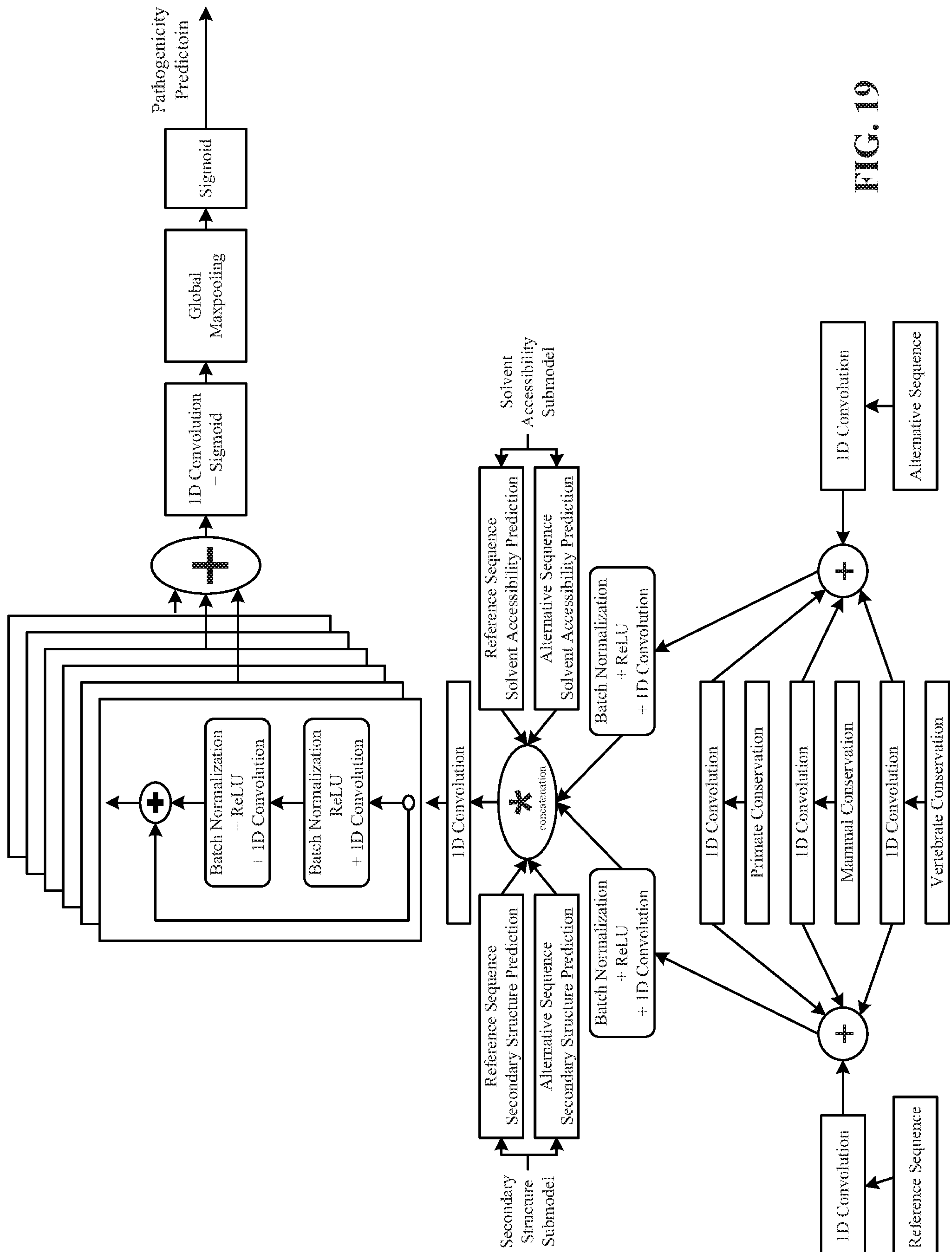


FIG. 18



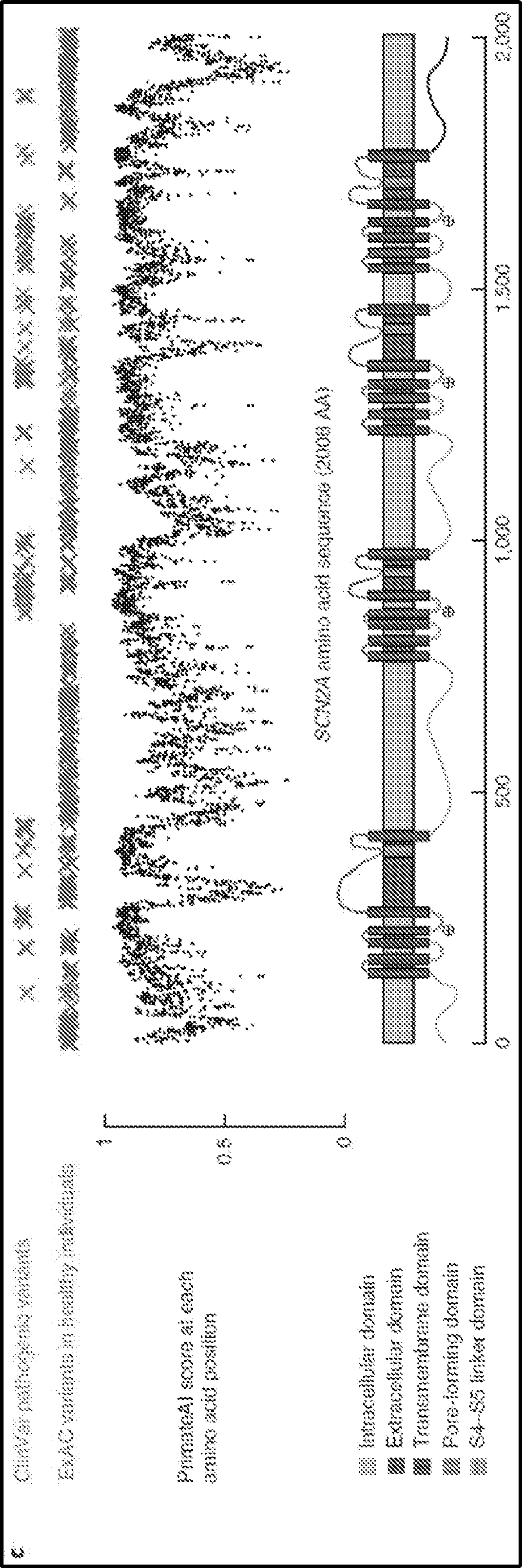
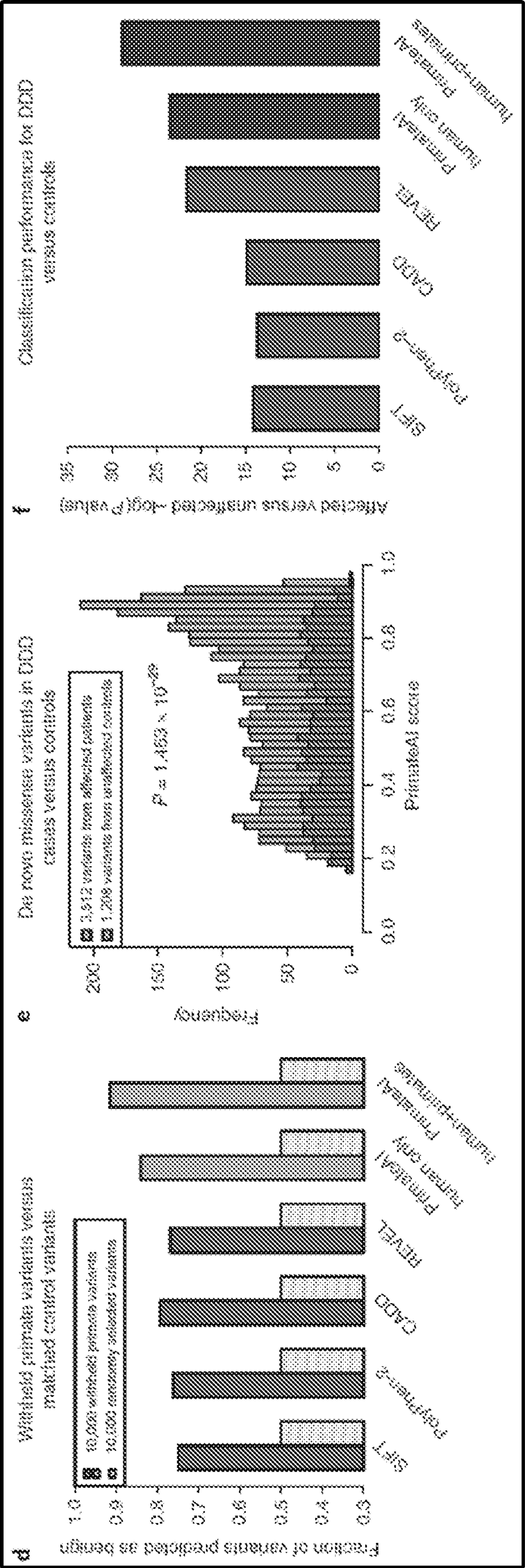
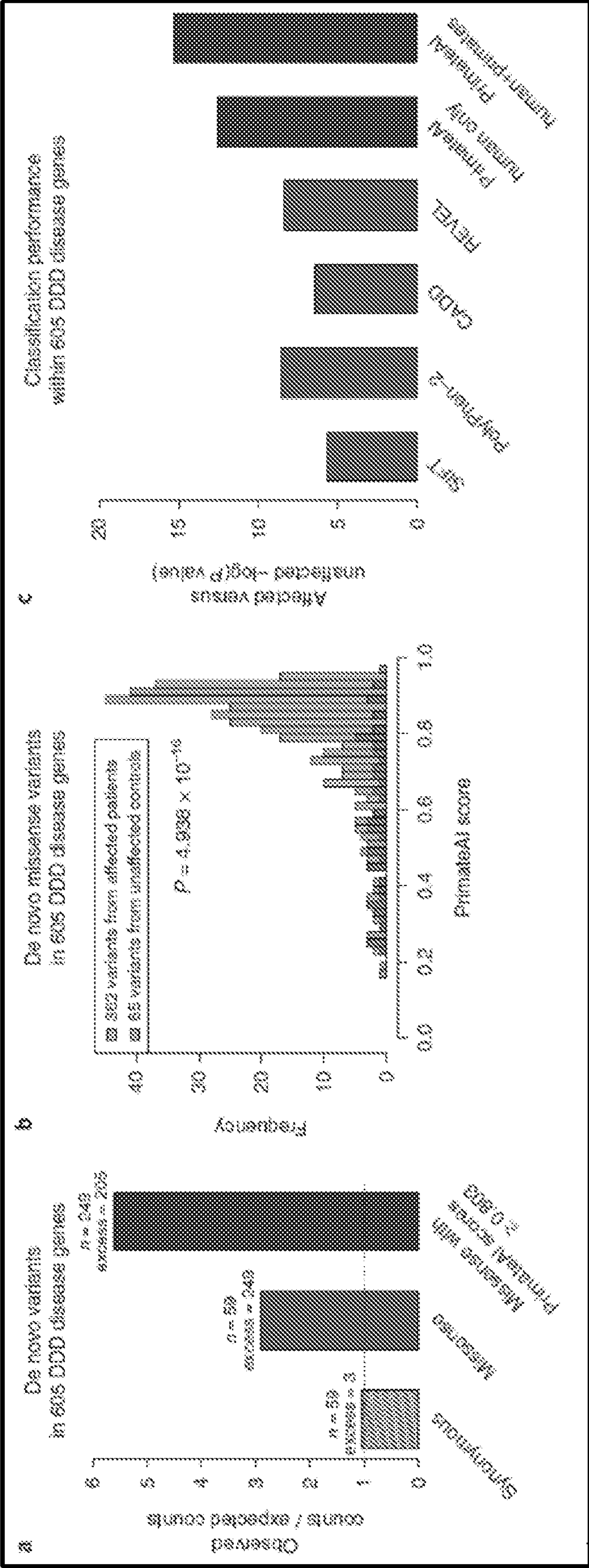


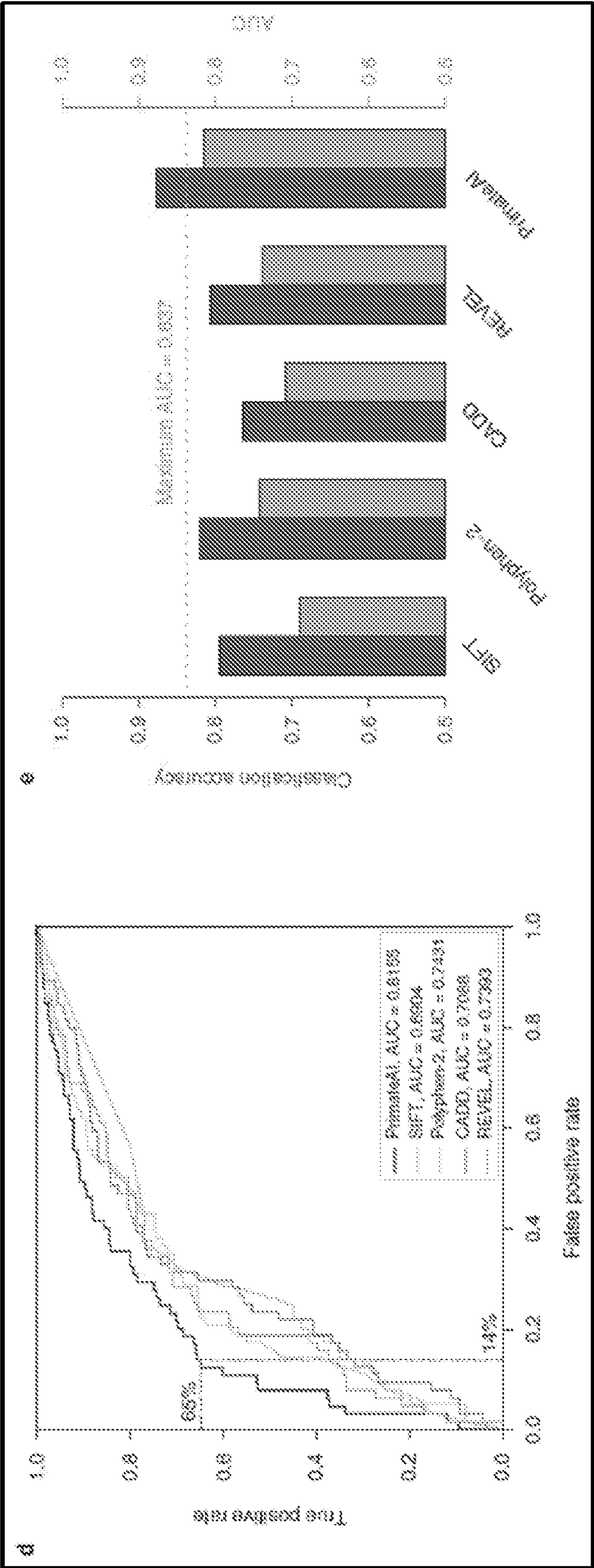
FIG. 20



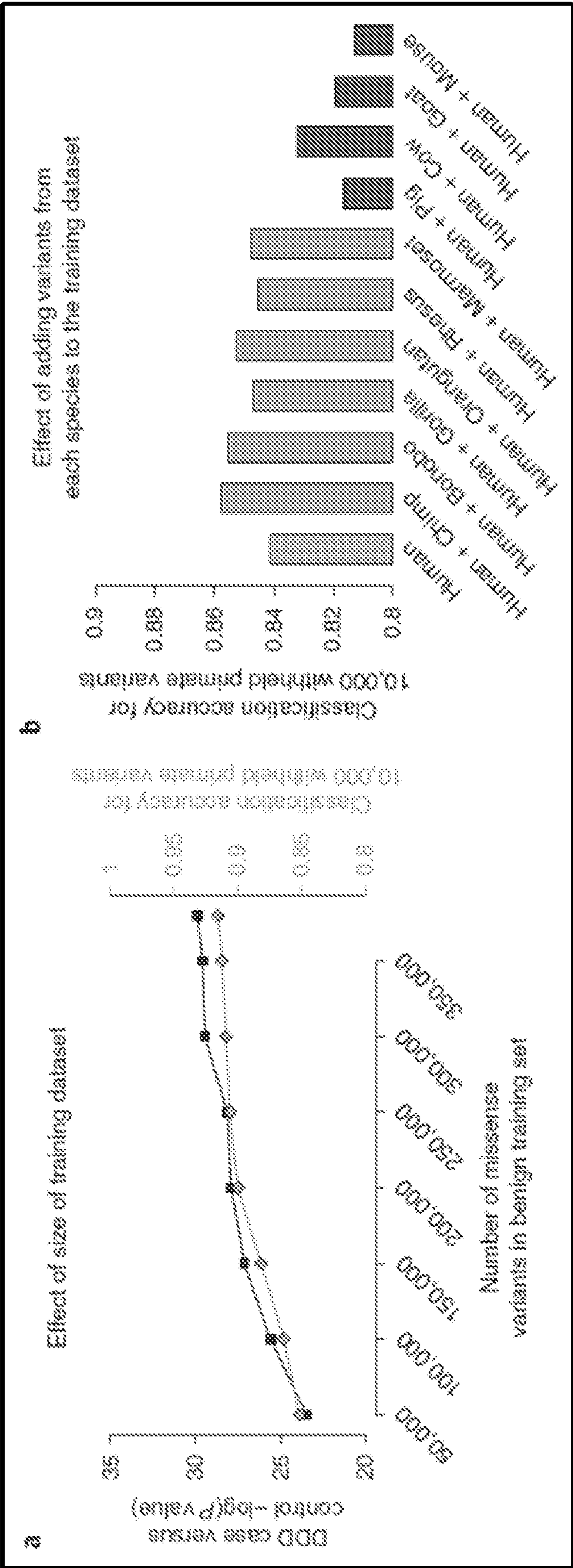
FIGs. 21D, 21E, 21F



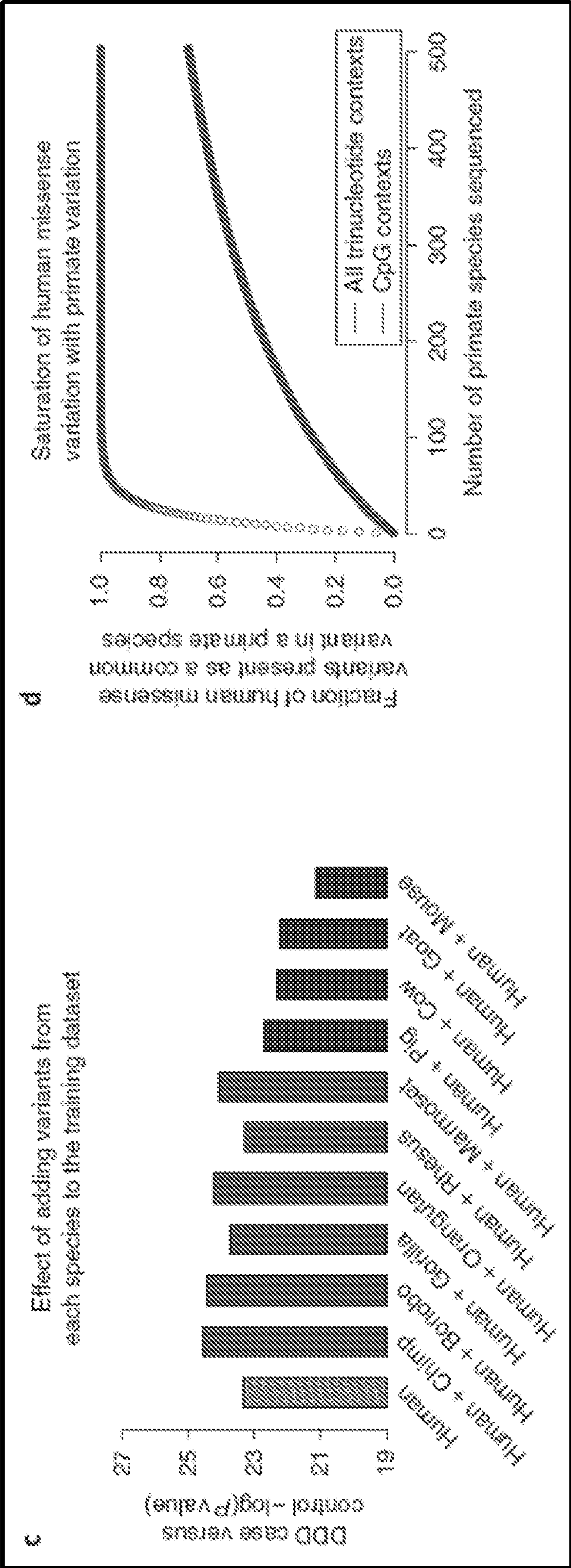
FIGs. 22A, 22B, 22C



FIGs. 22D, 22E



FIGs. 23A, 23B



FIGs. 23C, 23D

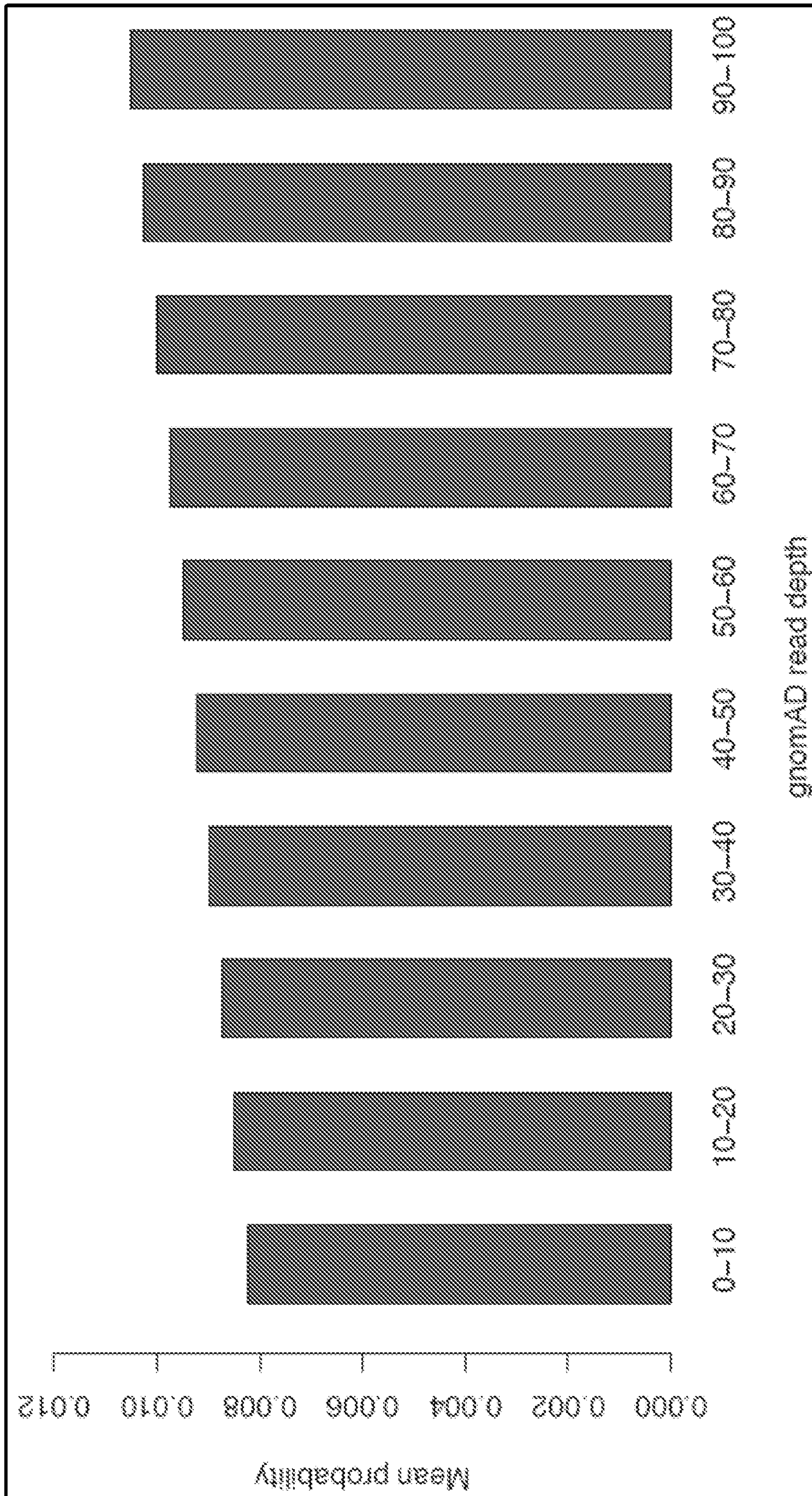


FIG. 24

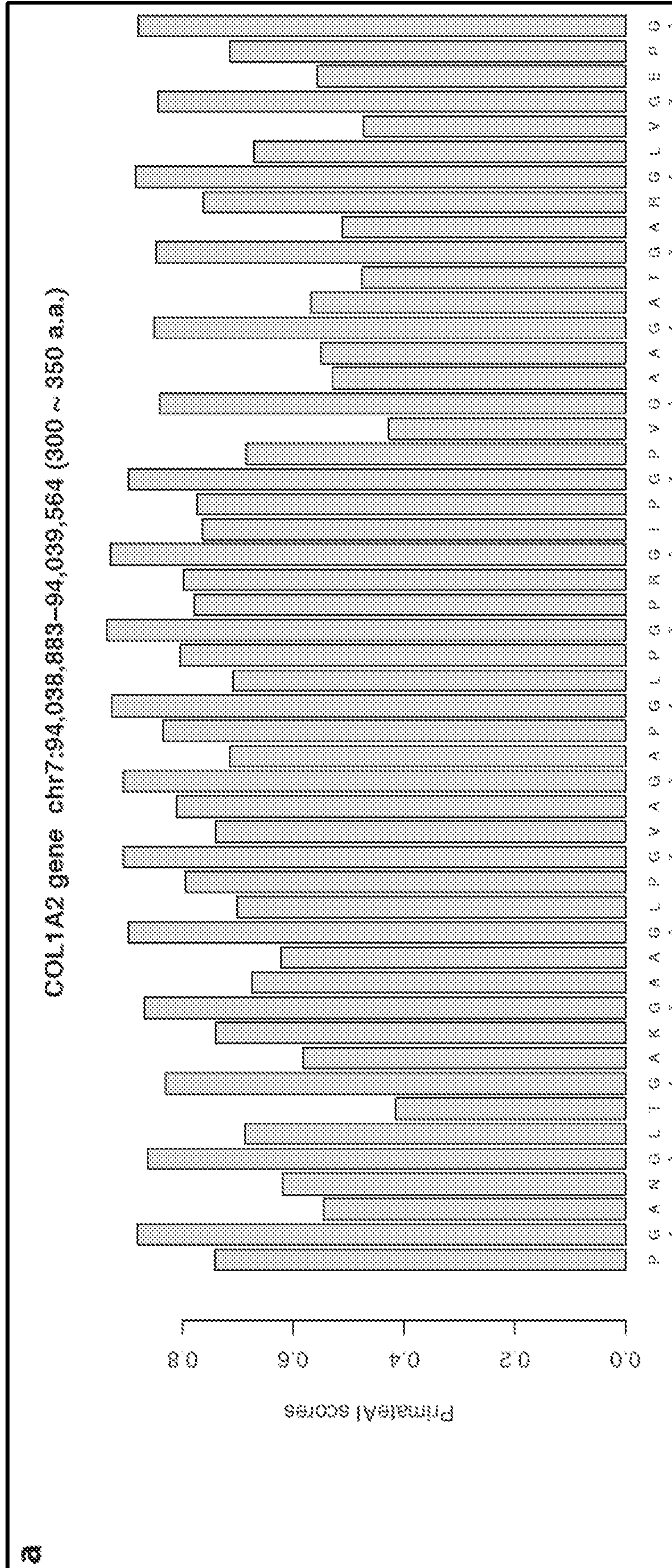


FIG. 25A

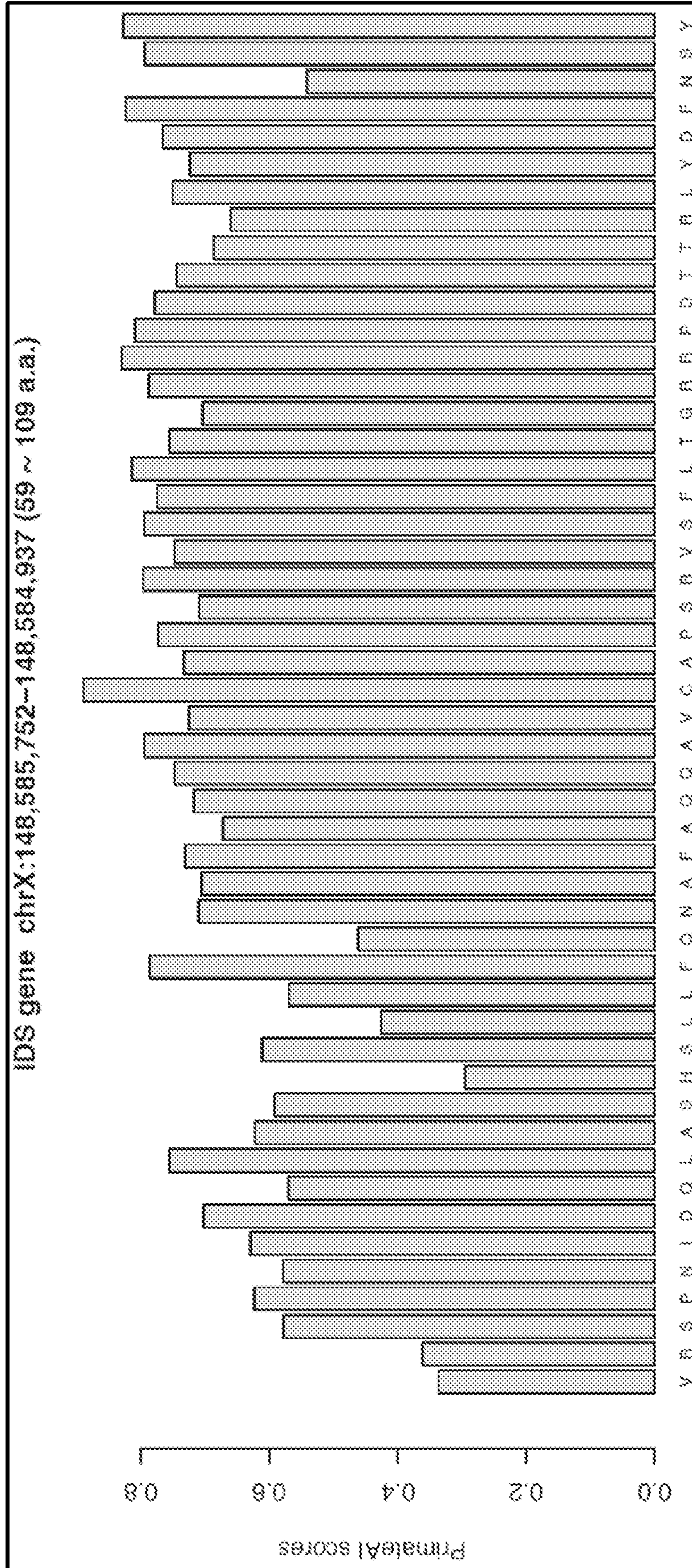


FIG. 25B

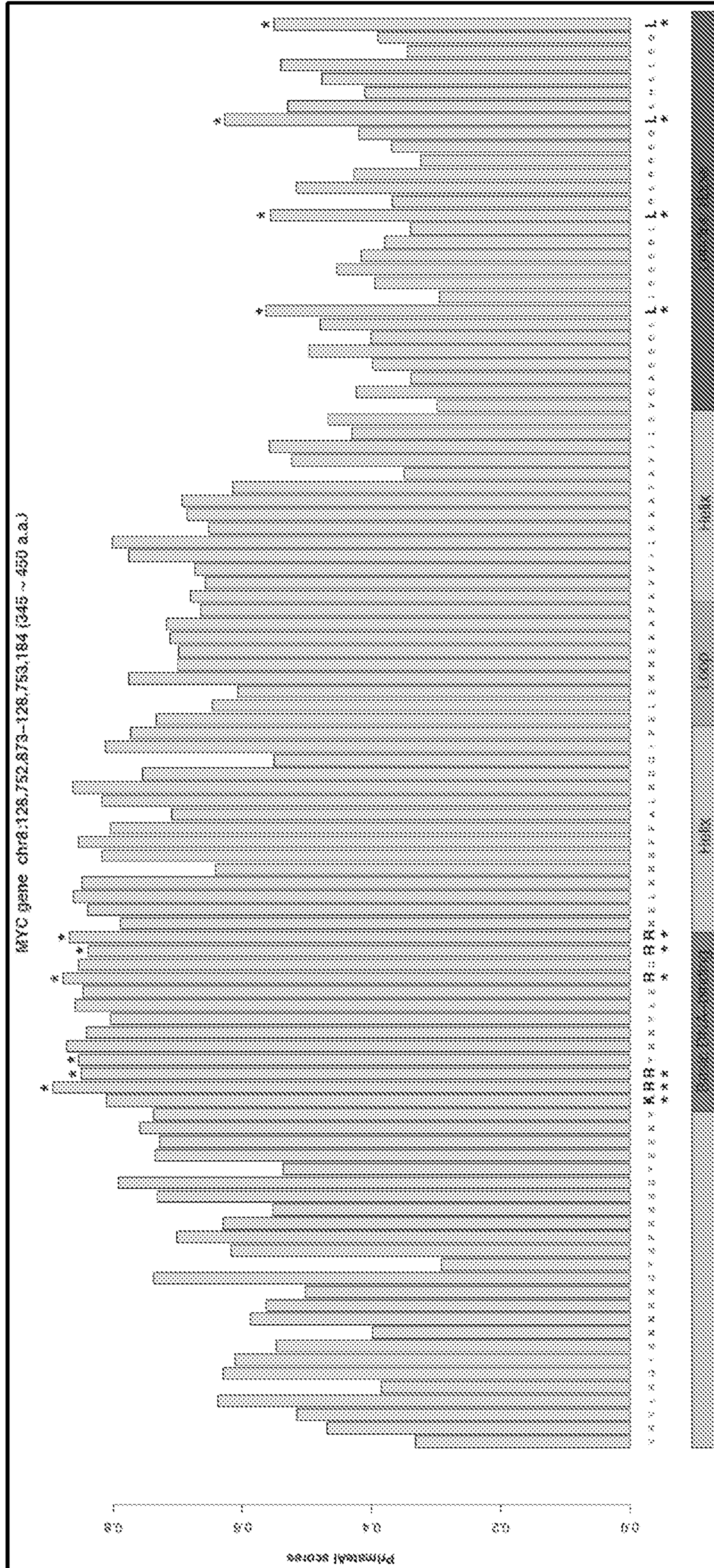


FIG. 25C

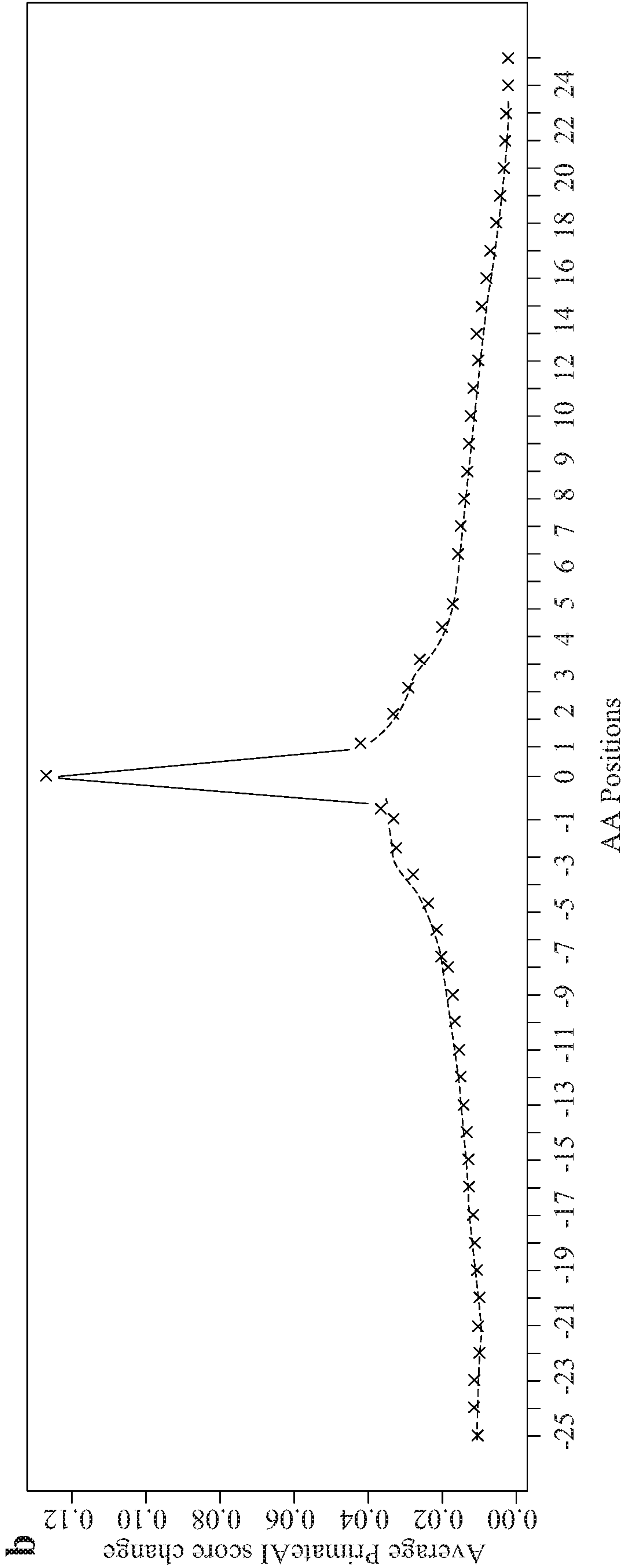


FIG. 26

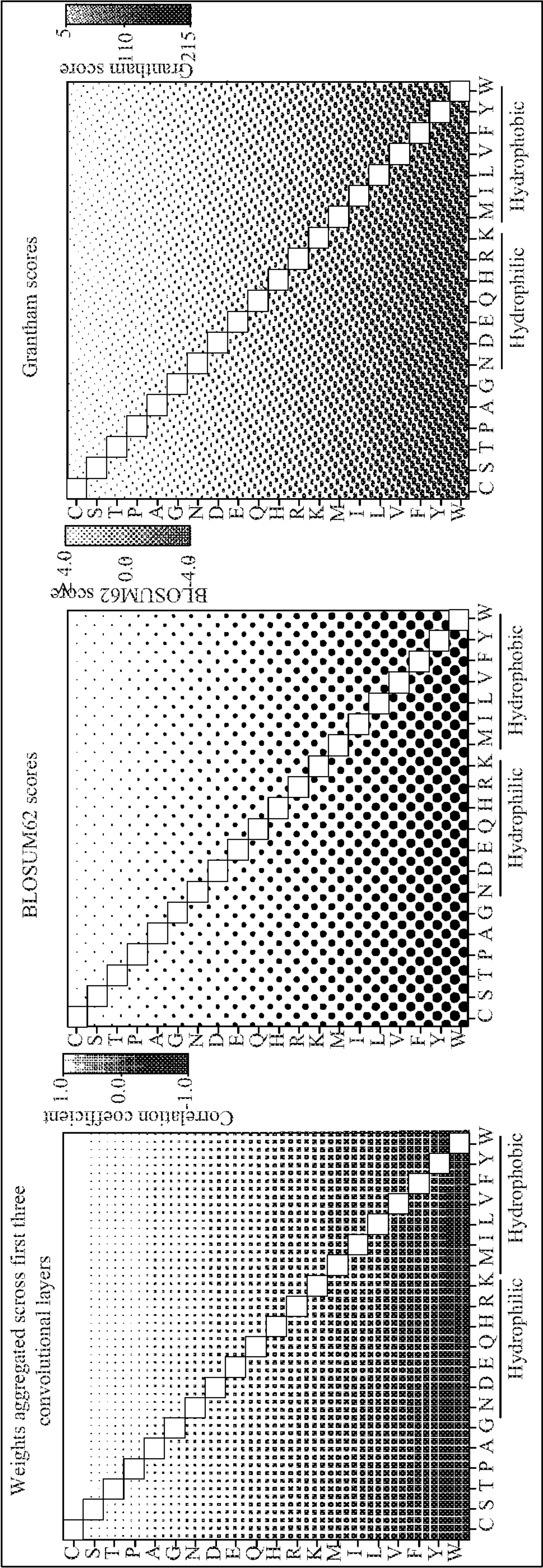


FIG. 27

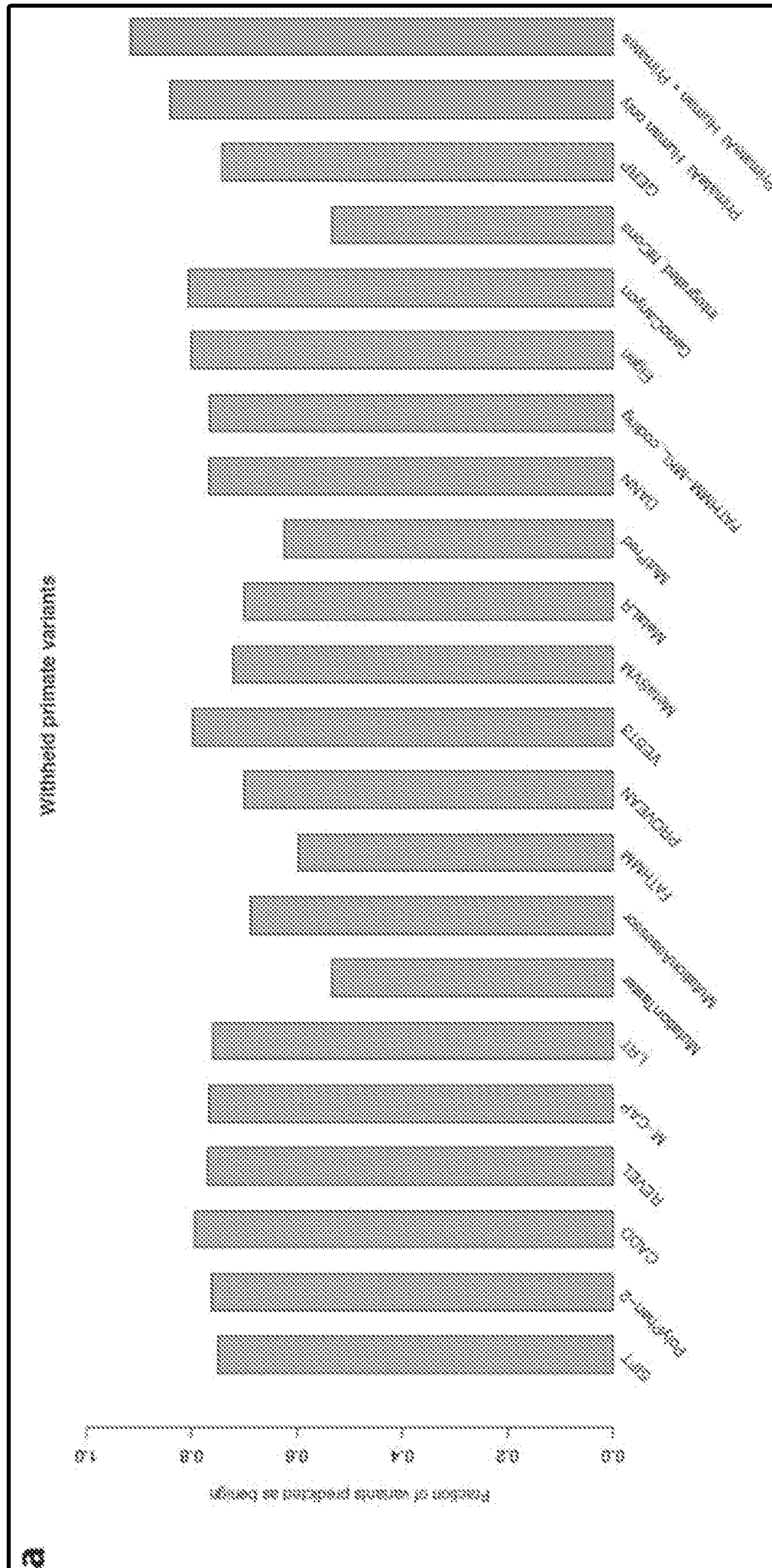


FIG. 28A

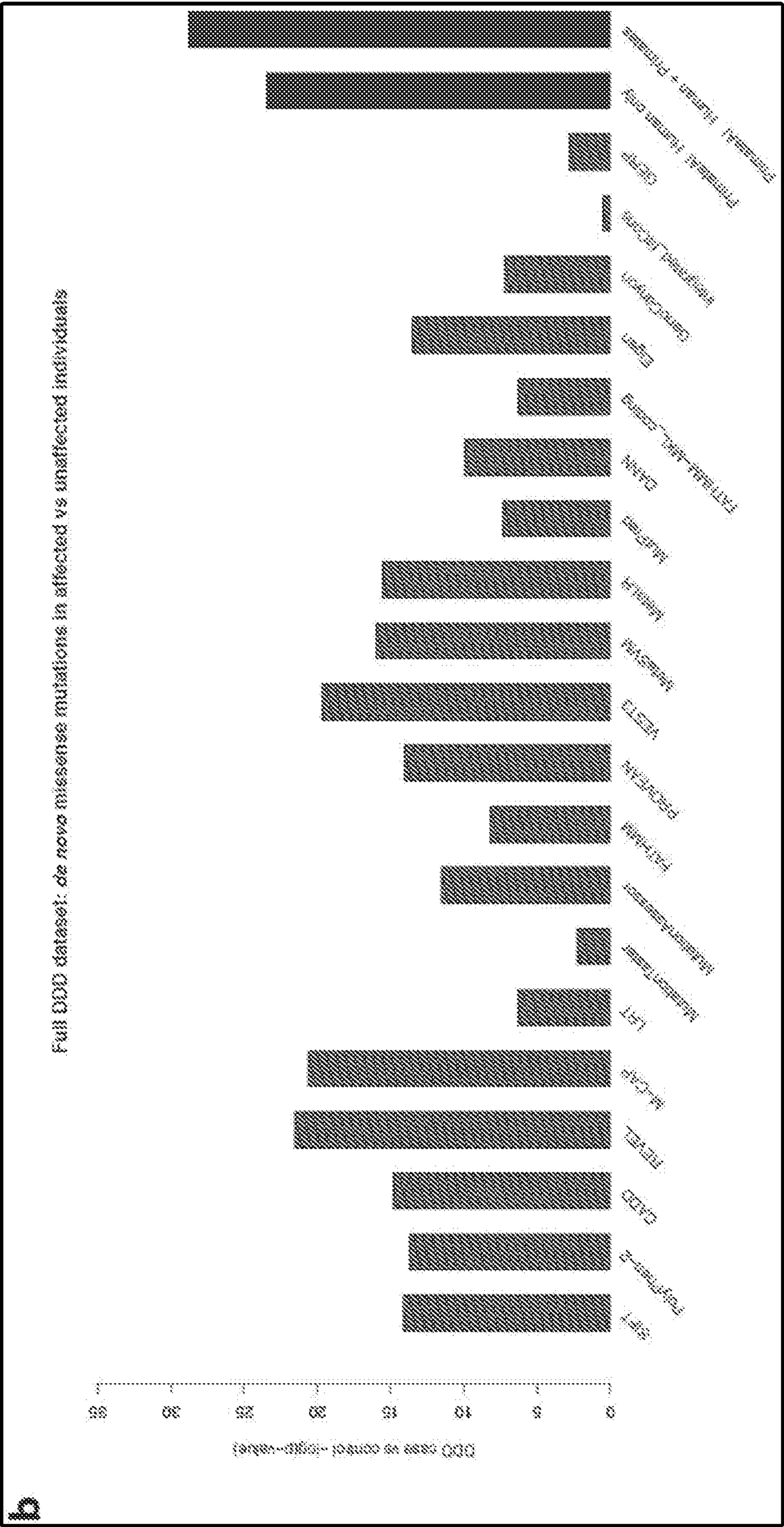


FIG. 28B

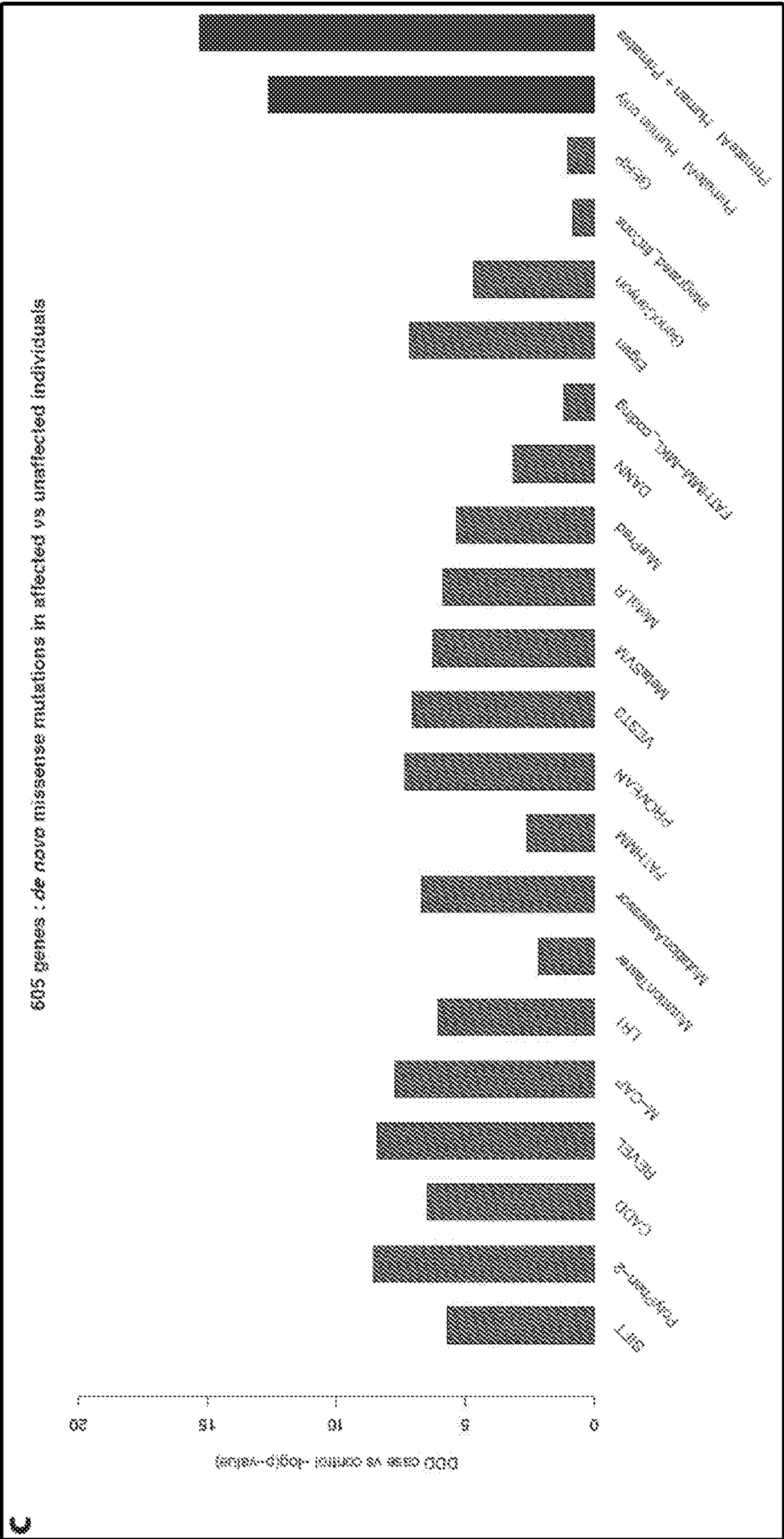


FIG. 28C

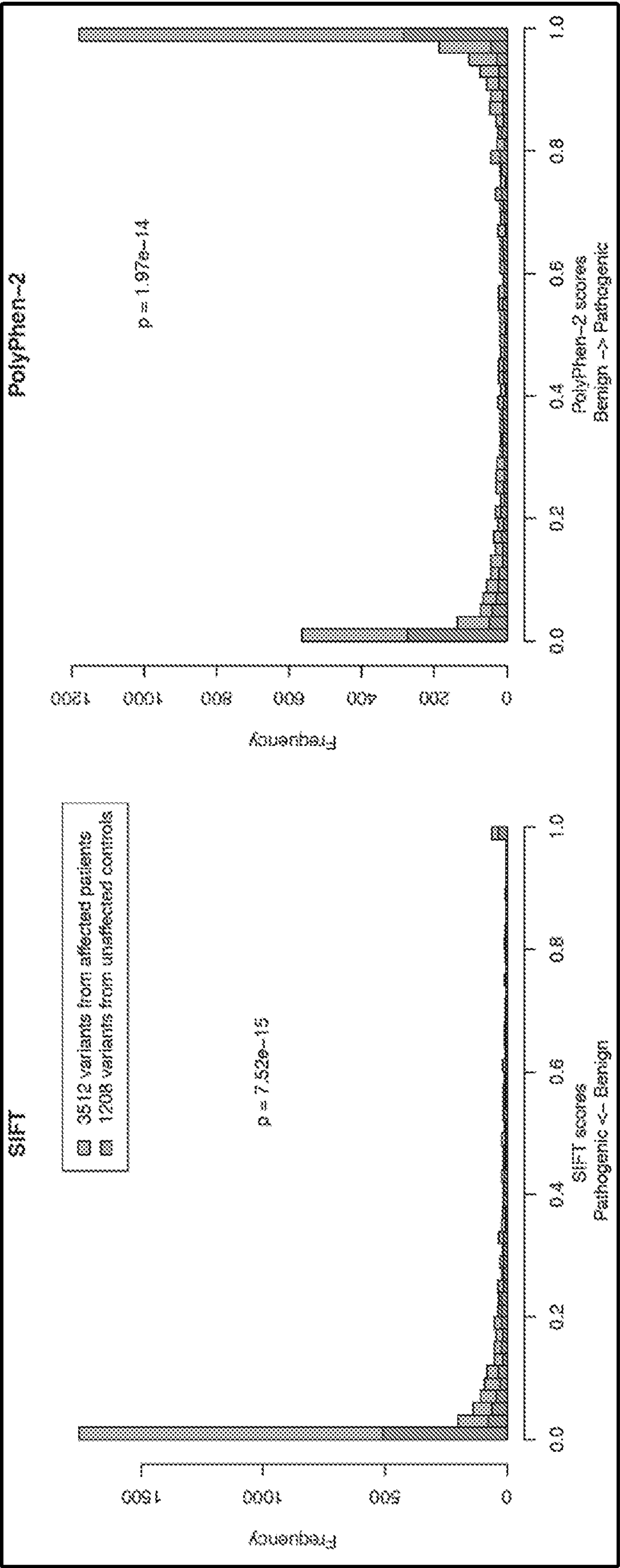


FIG. 29A

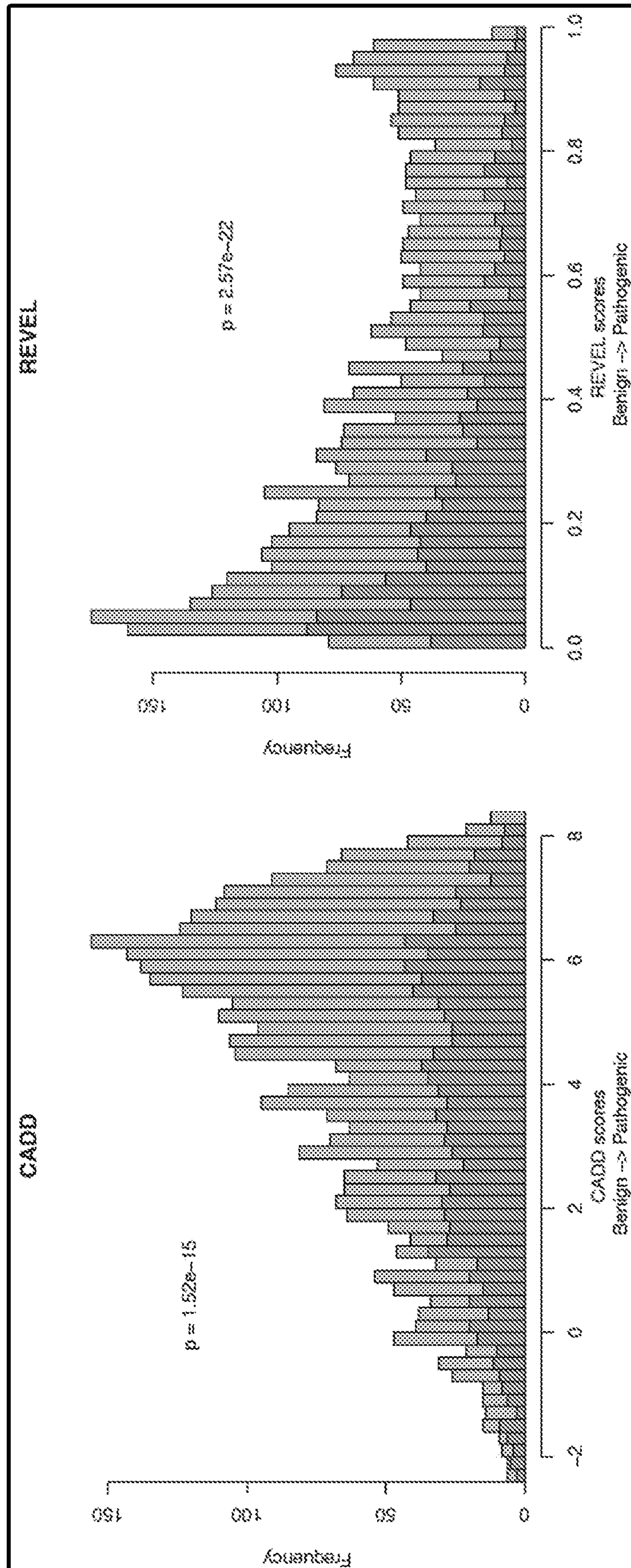


FIG. 29B

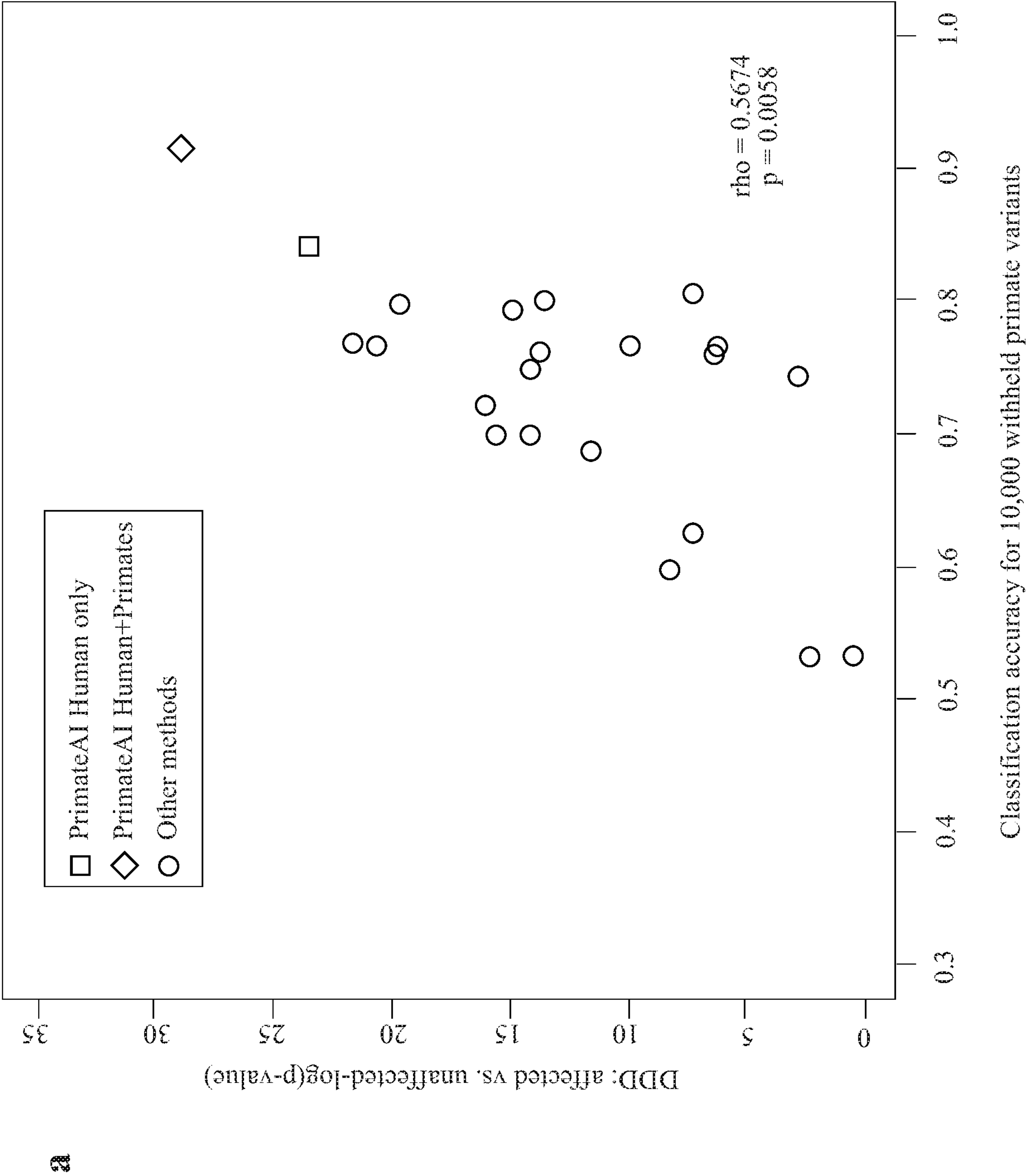


FIG. 30A

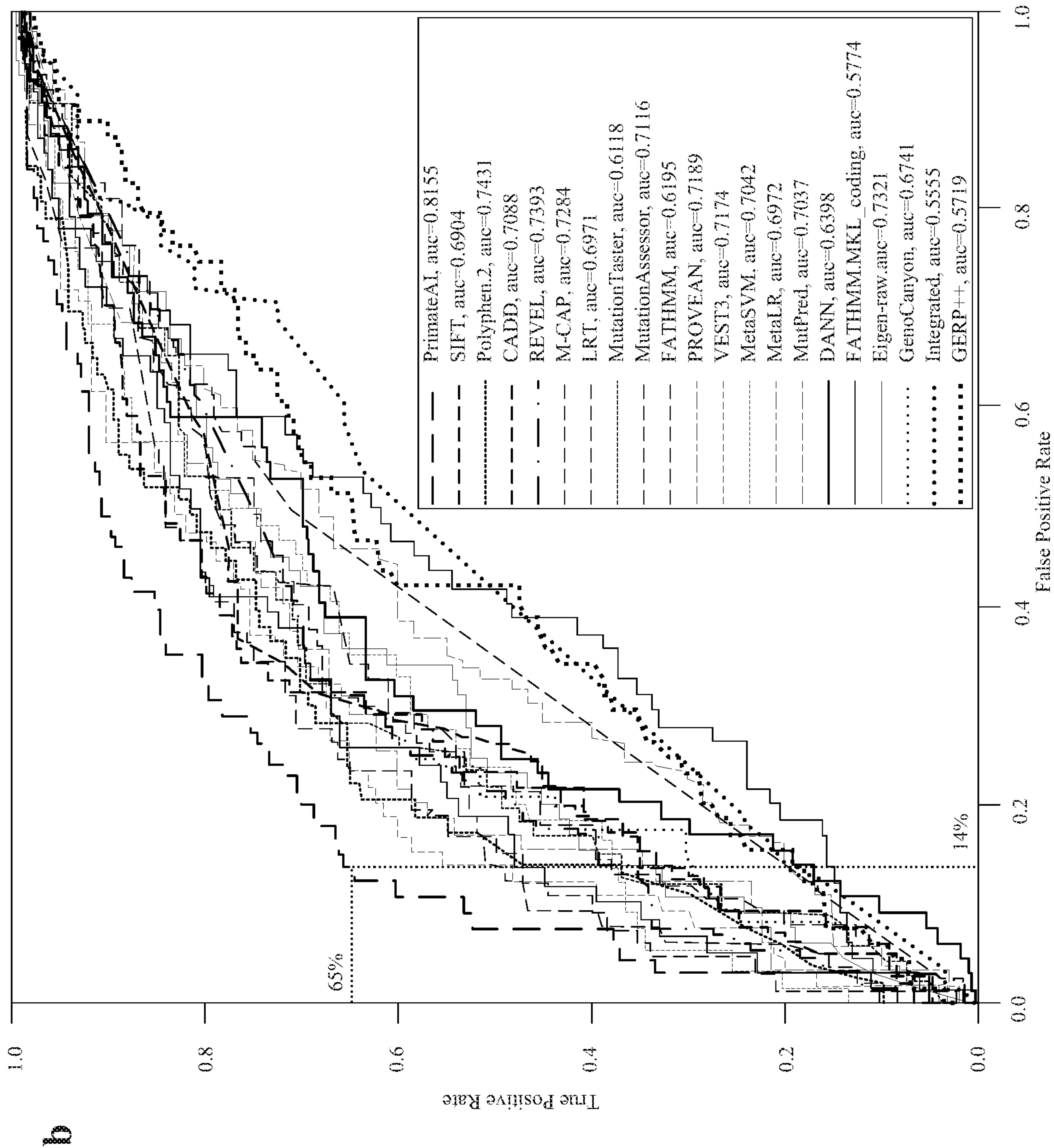


FIG. 30B

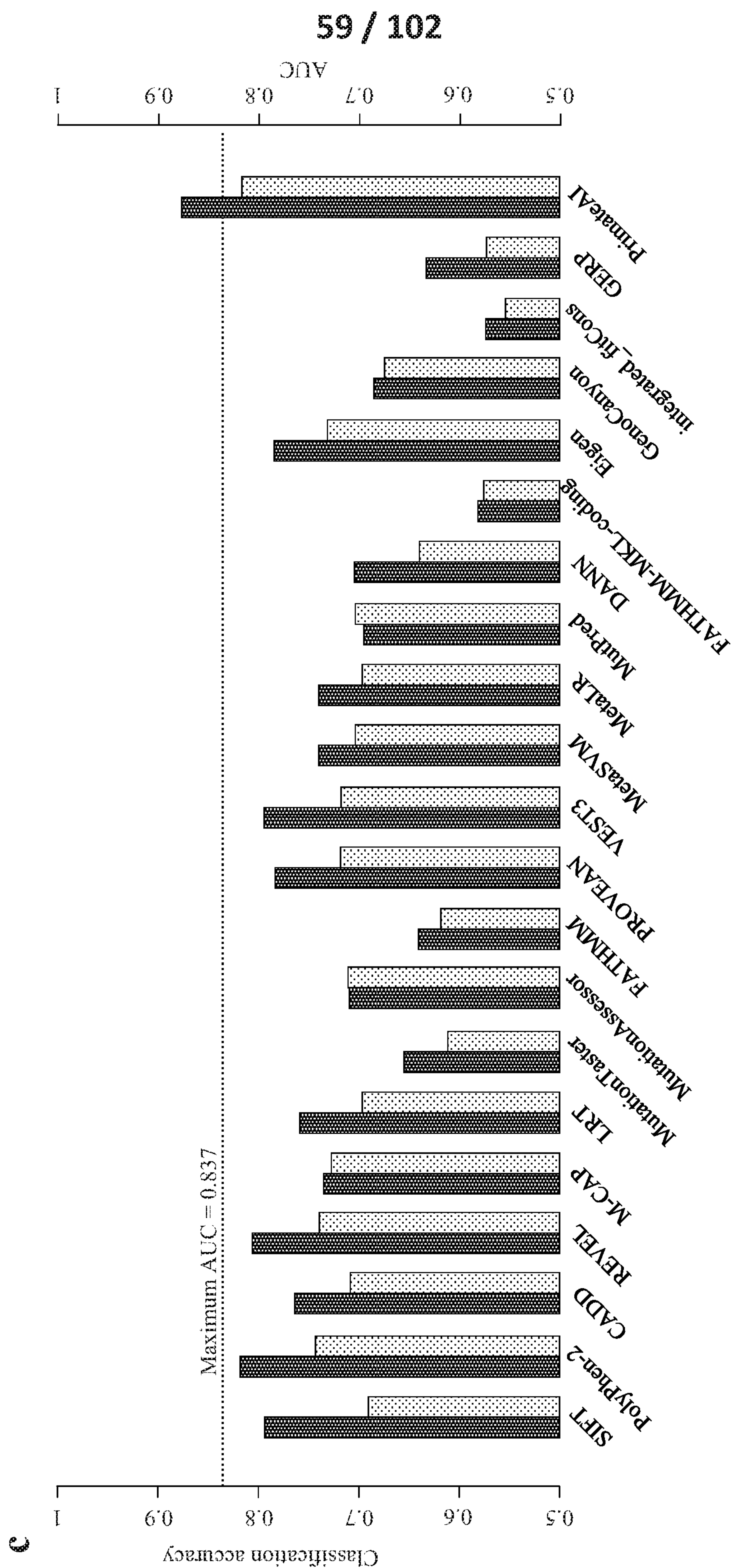


FIG. 30C

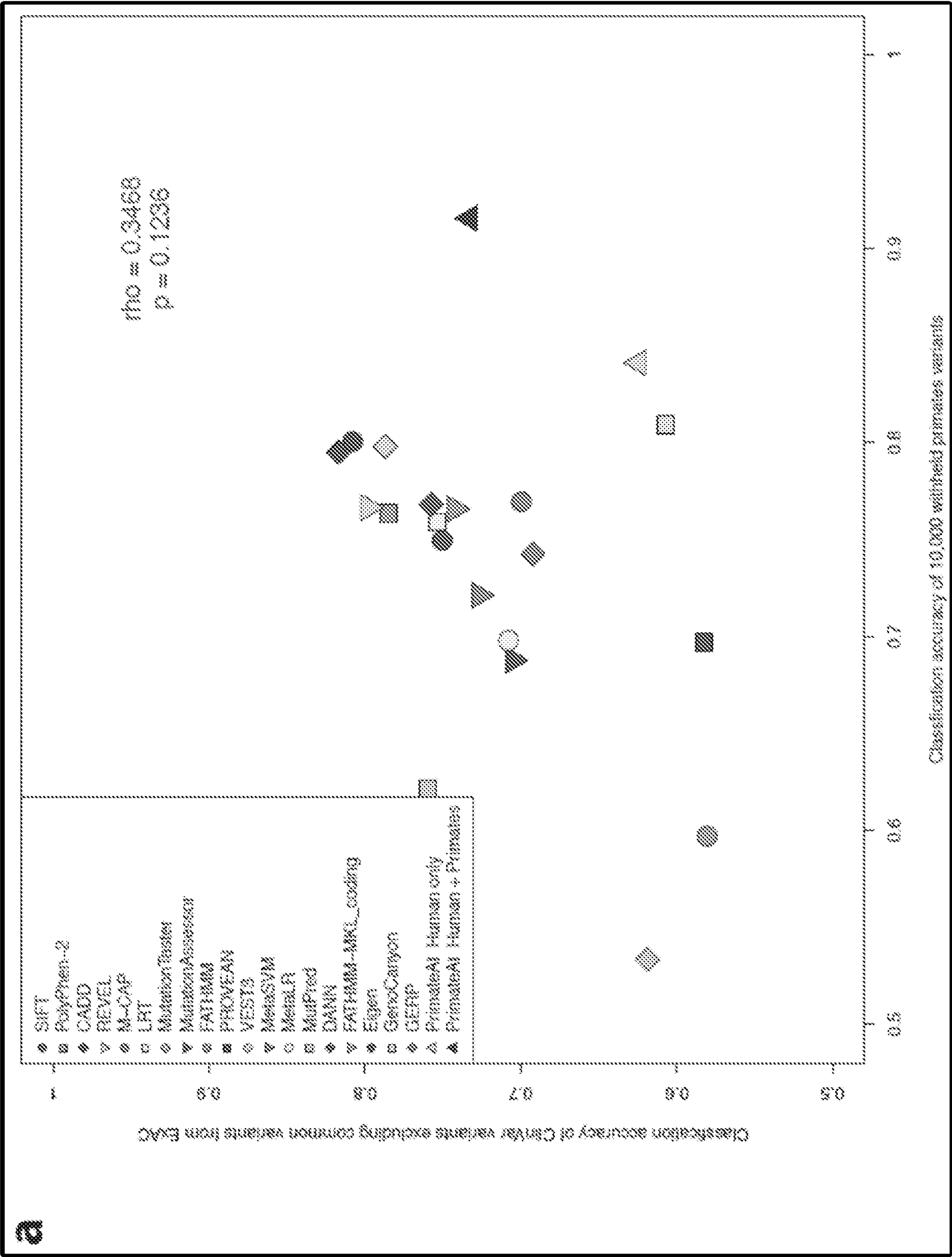


FIG. 31A

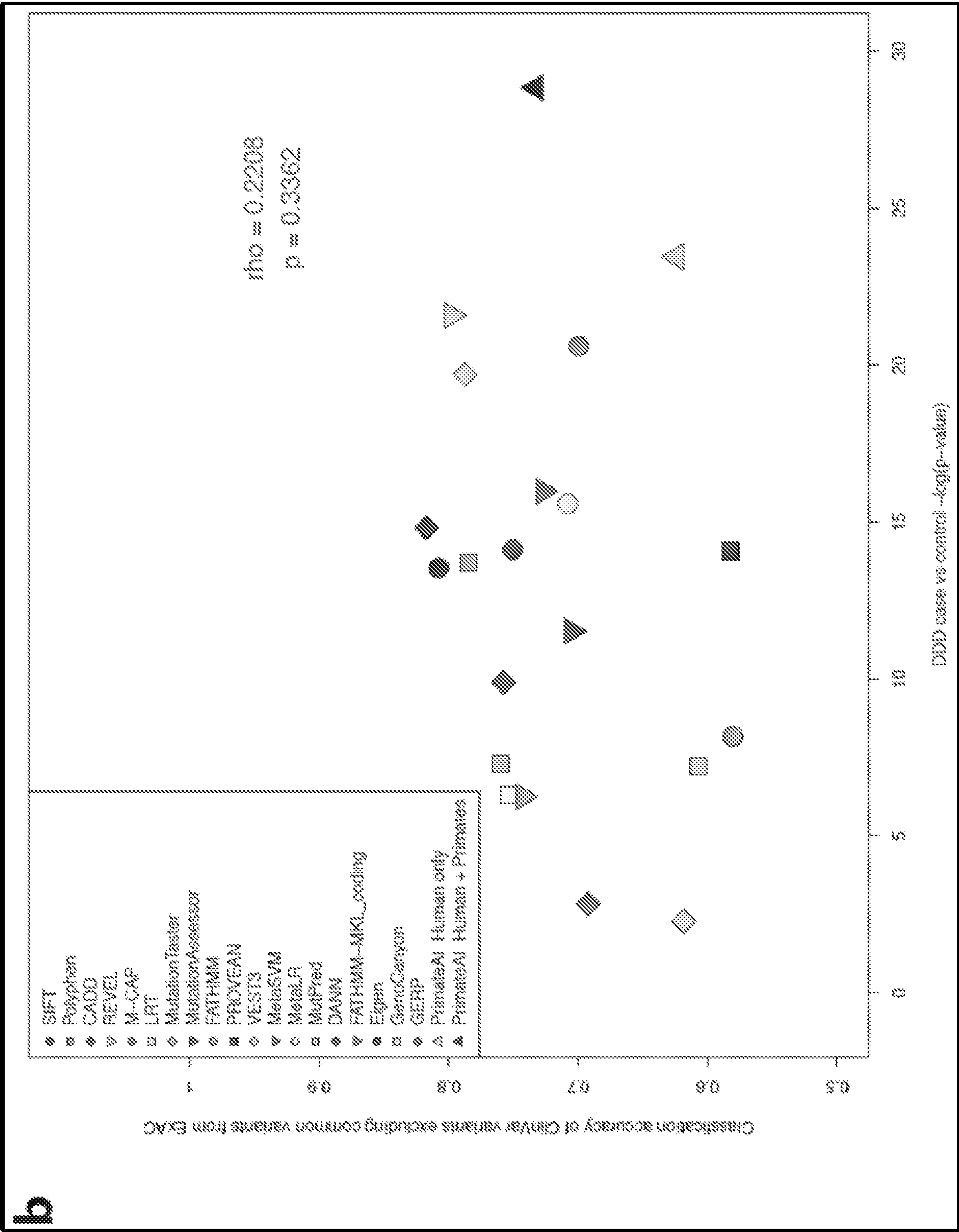


FIG. 31B

	Training Accuracy	Validation Accuracy	Testing Accuracy
3-State Secondary Structure	80.32 %	79.17 %	79.86 %
3-State Solvent Accessibility	64.83 %	60.53 %	60.31 %

FIG. 32

Models	Replicates	Withheld Primate Variant Accuracy	DDD Case vs Control -log(pvalue)	DDD 605 disease genes -log(pvalue)
Use DL predicted secondary structure labels	1	0.914	28.60	15.61
	2	0.919	32.32	15.94
	3	0.915	29.48	16.38
	4	0.917	30.35	15.62
	5	0.916	29.11	16.22
	Median	0.916	29.48	15.94
Replace DL predicted secondary structure labels with DSSP secondary structure labels of human proteins when available	1	0.916	29.48	15.94
	2	0.913	31.72	16.24
	3	0.915	31.09	15.69
	4	0.915	30.49	16.09
	5	0.915	30.57	14.79
	Median	0.915	30.79	16.09

FIG. 33

	Withheld Primate Variant Accuracy	DDD Case vs Control -log(pvalue)
SIFT	0.7490	14.1233
PolyPhen-2	0.7618	13.7044
CADD	0.7939	14.8167
M-CAP	0.7667	20.6078
LRT	0.7598	6.2960
MutationTaster	0.5332	2.2811
MutationAssessor	0.6879	11.5202
FATHMM	0.5981	8.1690
PROVEAN	0.6995	14.0747
VEST3	0.7978	19.7068
MetaSVM	0.7215	15.9829
MetaLR	0.6991	15.5842
REVEL	0.7686	21.5892
MutPred	0.6240	7.3044
DANN	0.7666	9.8976
MKL_coding	0.7657	6.2469
Eigen	0.8003	13.5364
GenoCanyon	0.8058	7.2196
integrated_fitCons	0.5331	0.4954
GERP	0.7430	2.8285
PrimateAI Human	0.8411	23.4536
PrimateAI Human + Primates	0.9156	28.8346

FIG. 34

	-log(pvalue)	AUC	Threshold	Classification	
				Accuracy	
SIFT	5.6793	0.6903	0.001	0.7949	
PolyPhen-2	8.5522	0.7430	0.956	0.8197	
CADD	6.4772	0.7087	5.040	0.7643	
M-CAP	7.7143	0.7283	0.081	0.7348	
LRT	6.0285	0.6970	1.00E-06	0.7604	
MutationTaster	2.1366	0.6117	1.000	0.6564	
MutationAssessor	6.6779	0.7115	2.240	0.7096	
FATHMM	2.5979	0.6194	-1.260	0.6401	
PROVEAN	7.3447	0.7188	-4.310	0.7833	
VEST3	7.0357	0.7173	0.679	0.7953	
MetaSVM	6.2556	0.7041	-0.3371	0.7406	
MetaLR	5.8707	0.6971	0.349	0.7407	
REVEL	8.4115	0.7392	0.456	0.8056	
MutPred	5.3370	0.7036	0.526	0.6949	
DANN	3.1588	0.6397	0.996	0.7046	
MKL_coding	1.1649	0.5773	0.965	0.5840	
Eigen	7.1659	0.7320	0.577	0.7844	
GenoCanyon	4.6763	0.6740	0.999	0.6875	
integrated_fitCons	0.8315	0.5554	0.706	0.5736	
GERP	1.0330	0.5718	5.060	0.6330	
PrimateAI Human					
+ Primates	15.3064	0.8154	0.802	0.8772	

FIG. 35

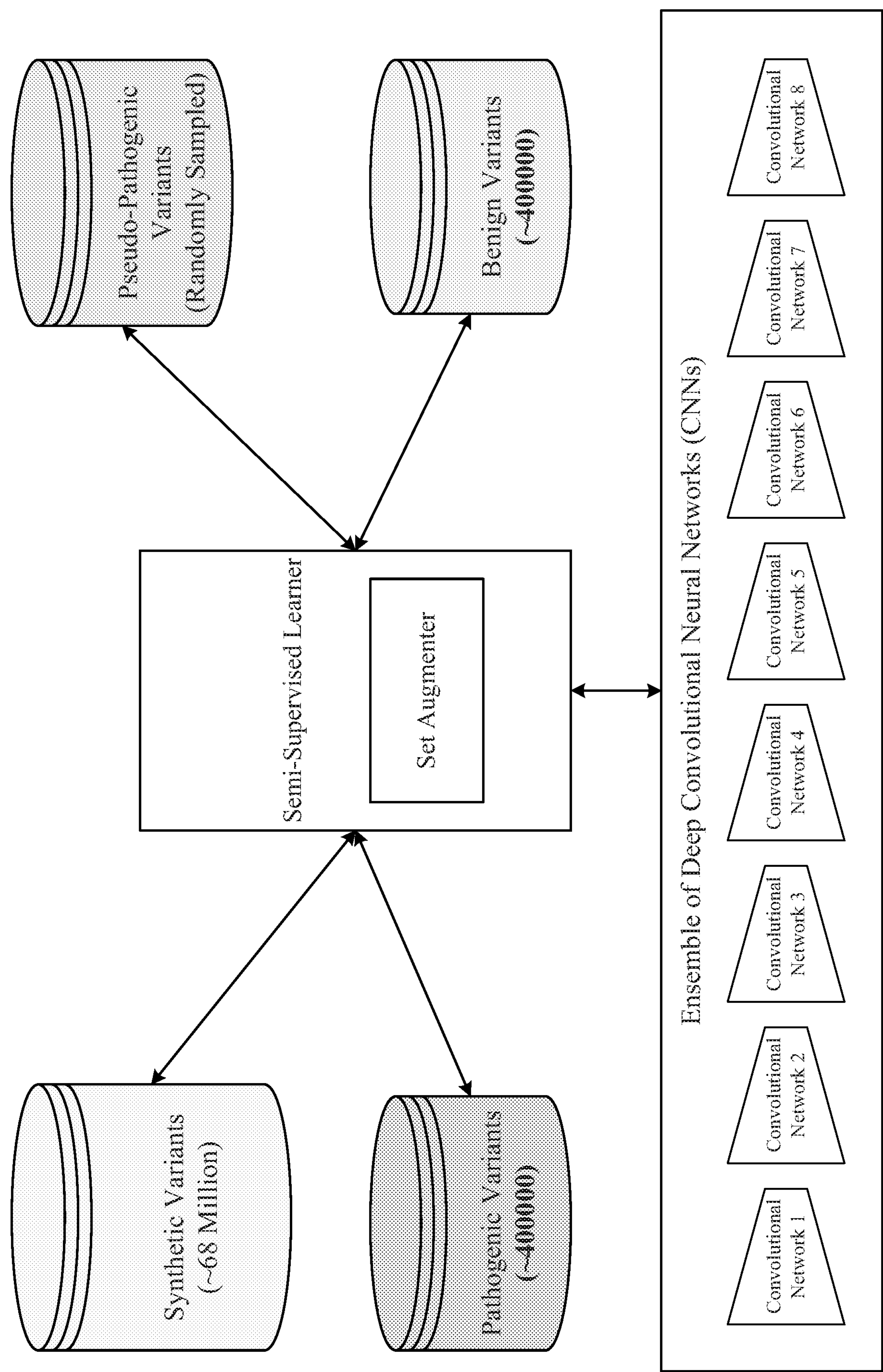


FIG. 36

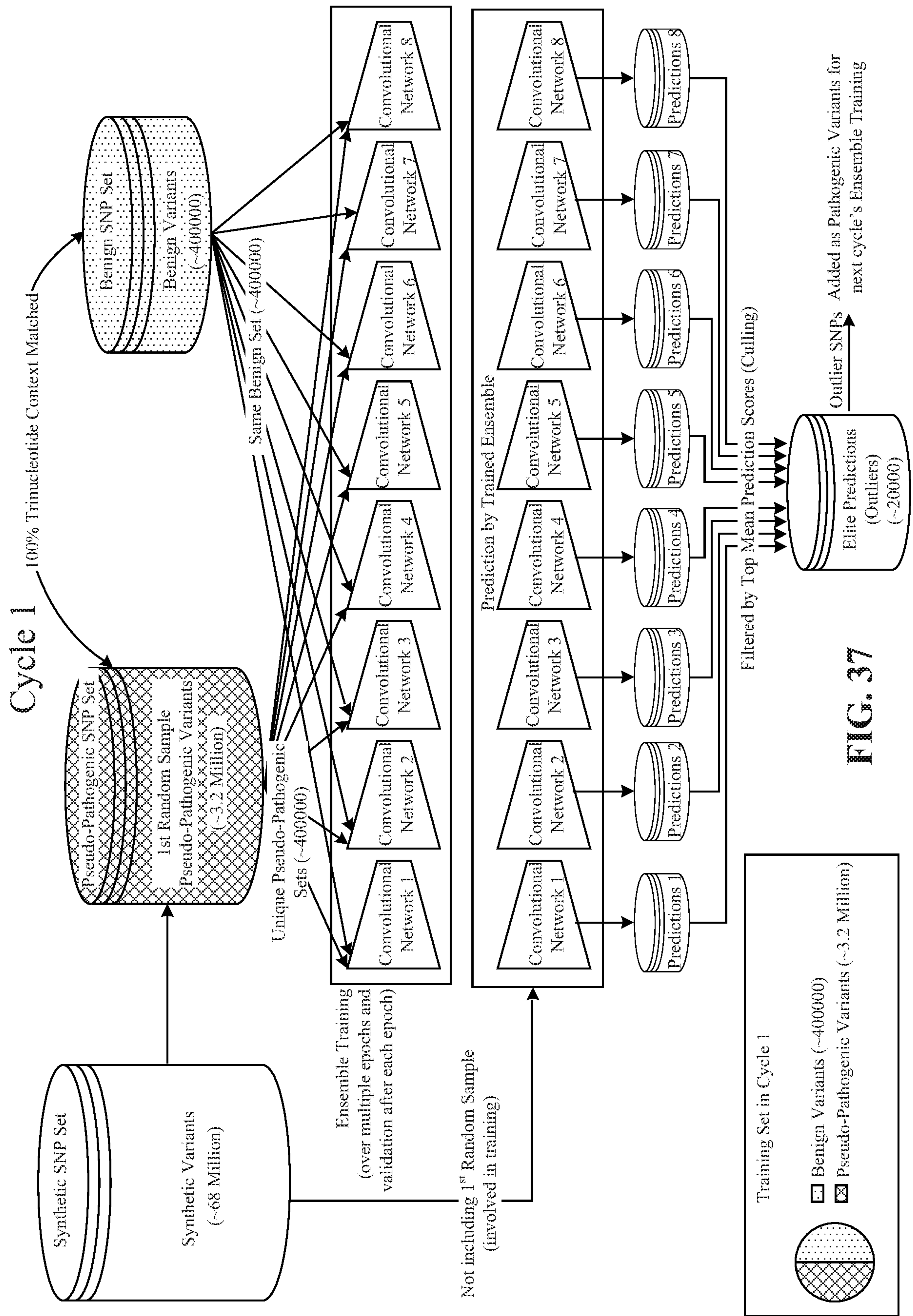


FIG. 37



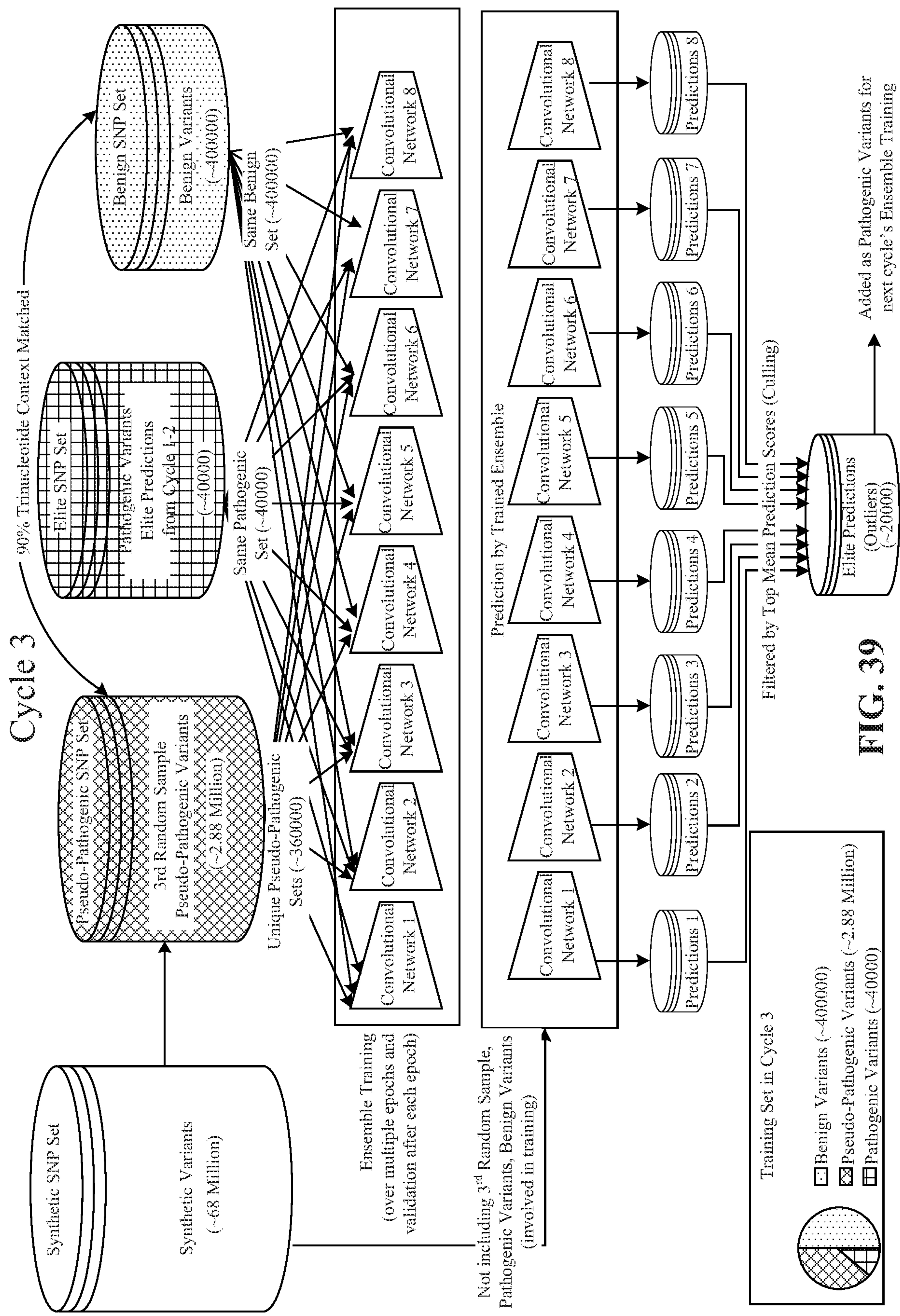


FIG. 39

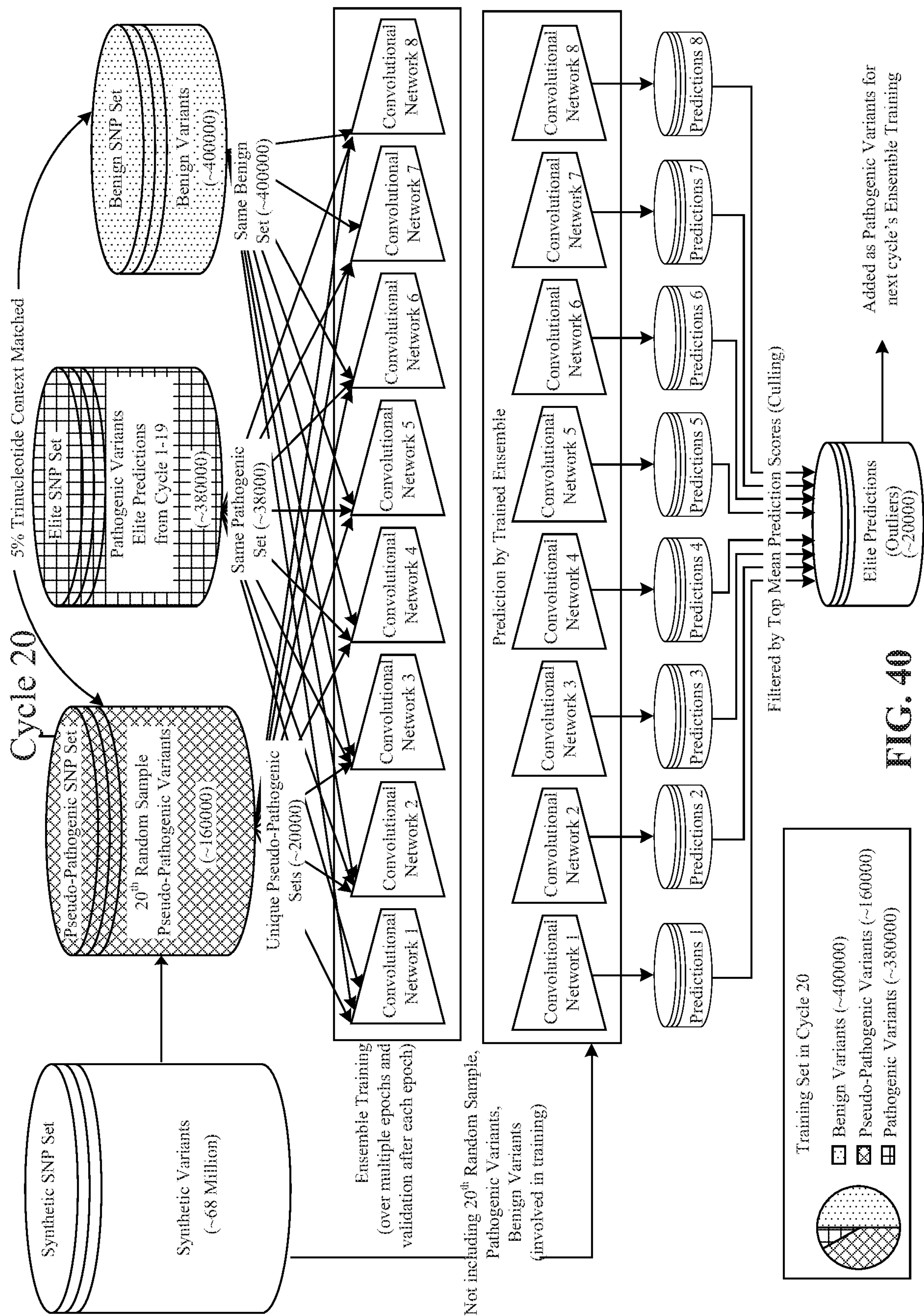


FIG. 40

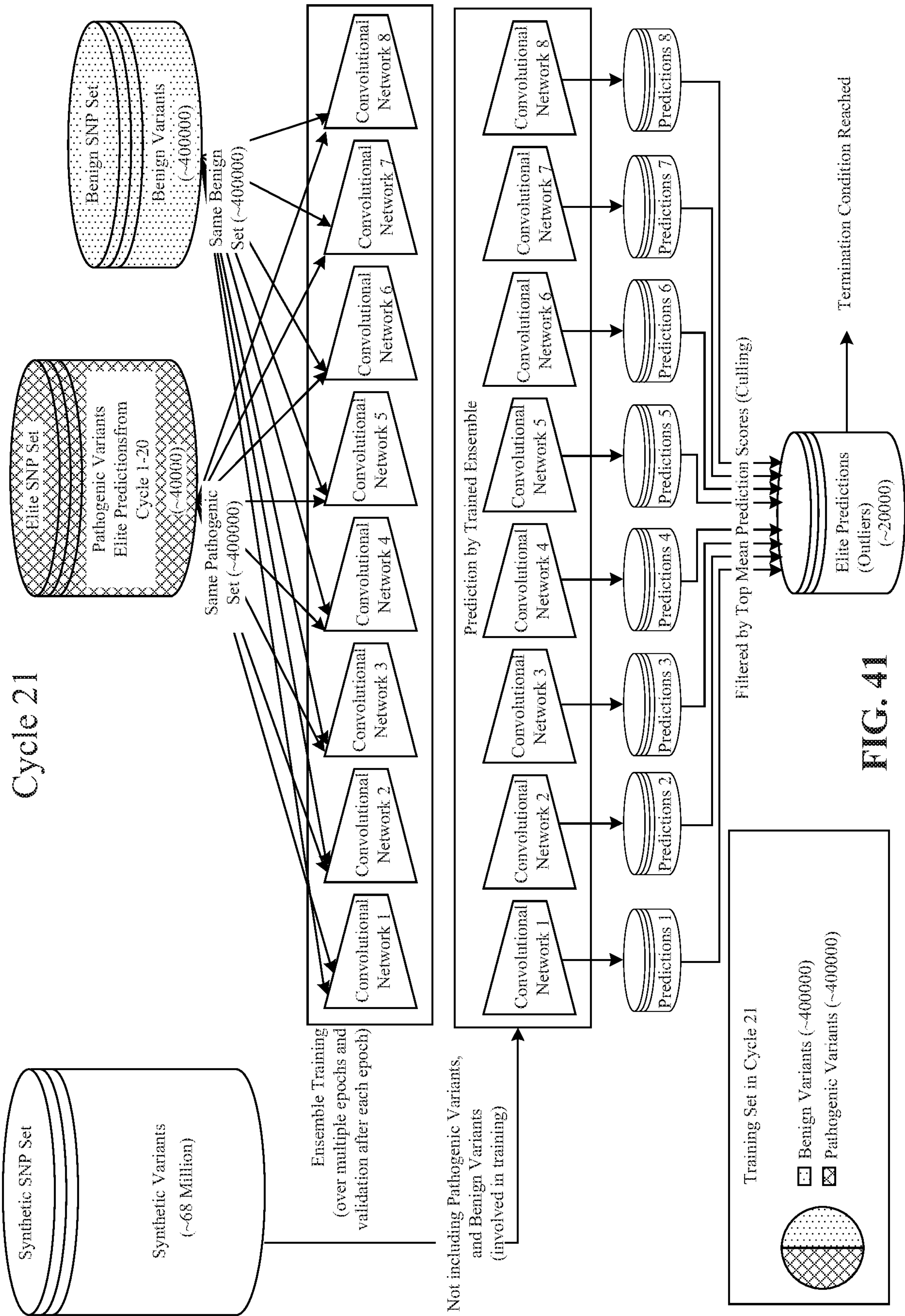


FIG. 41

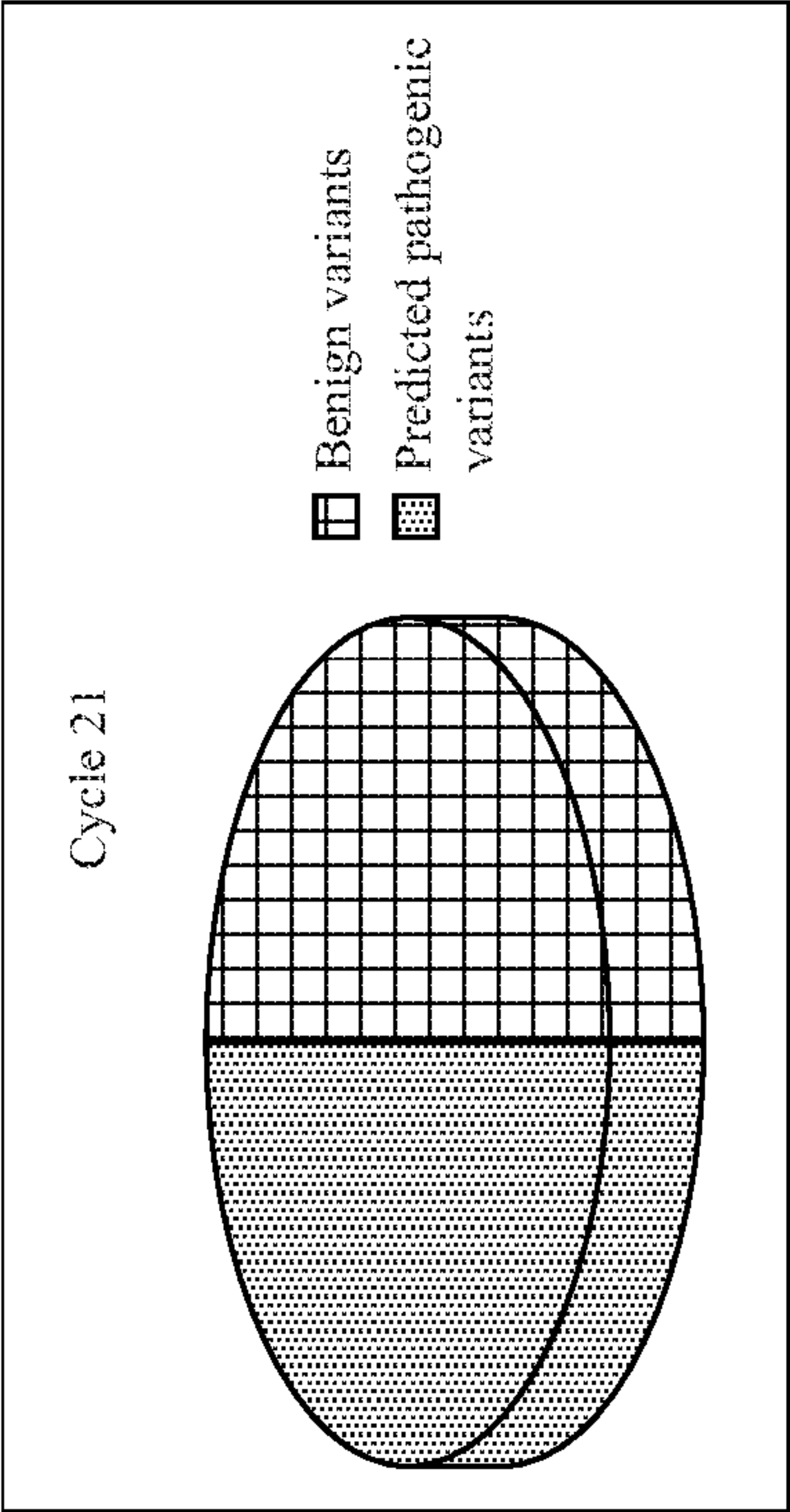
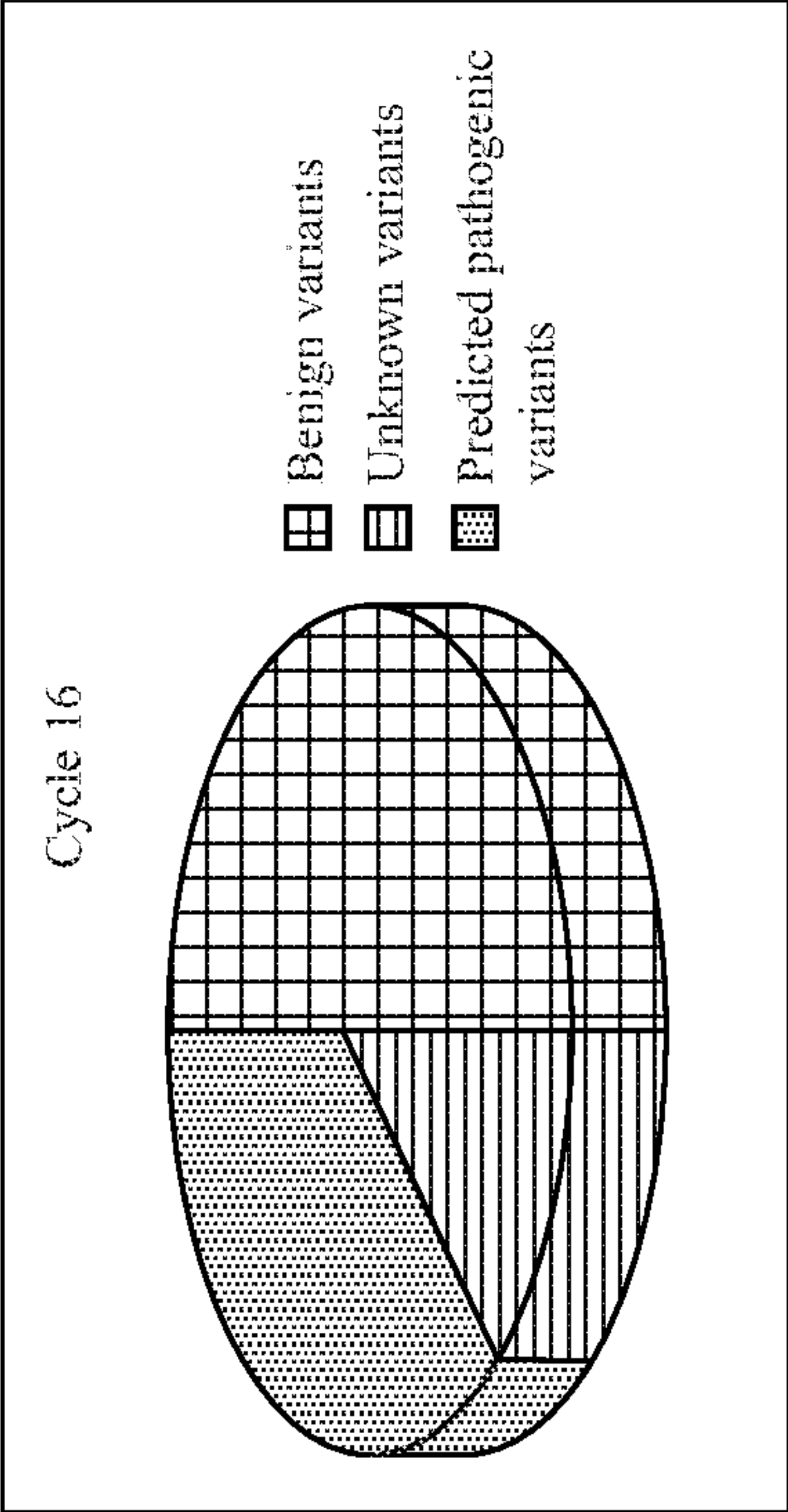
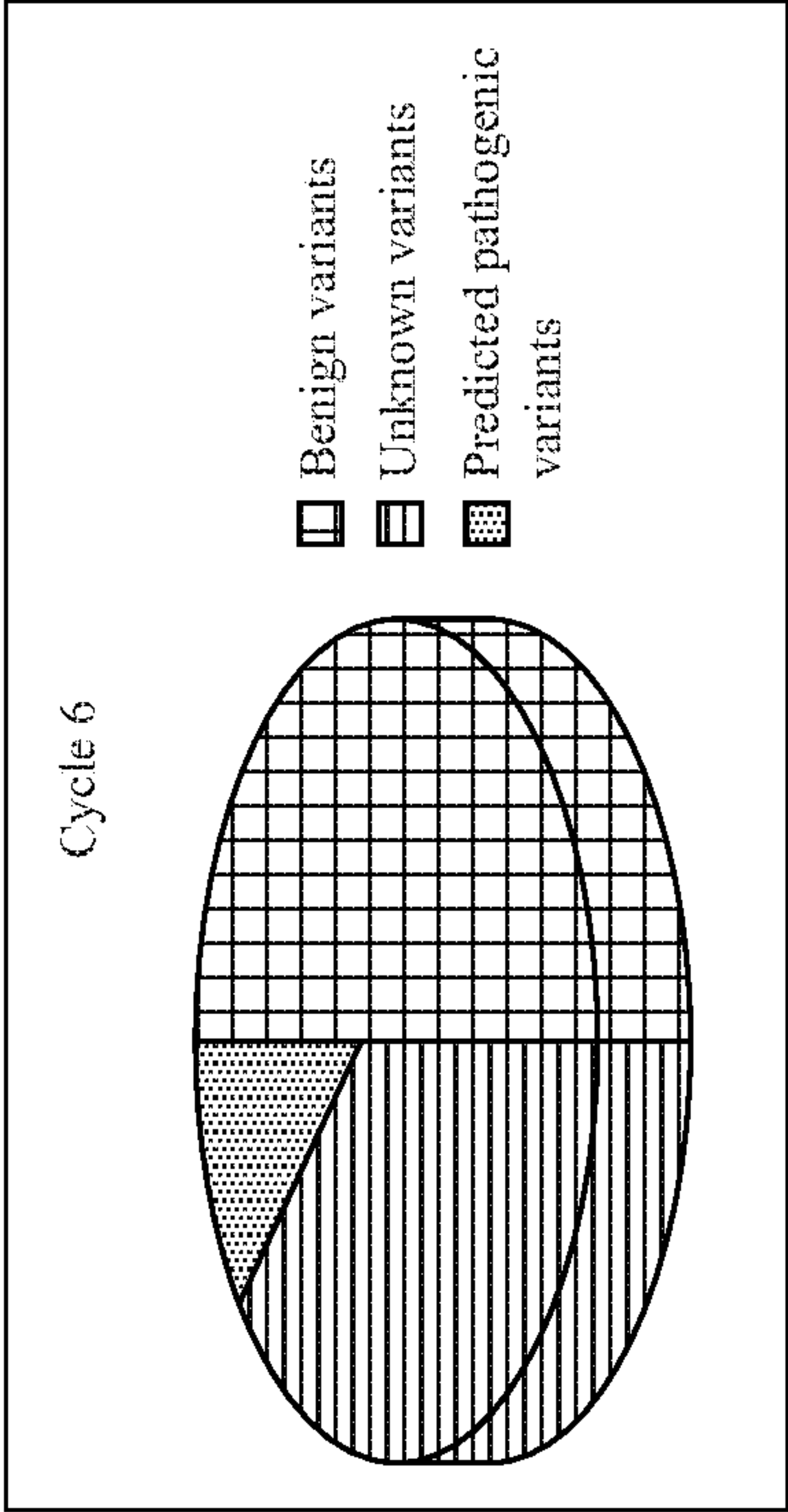
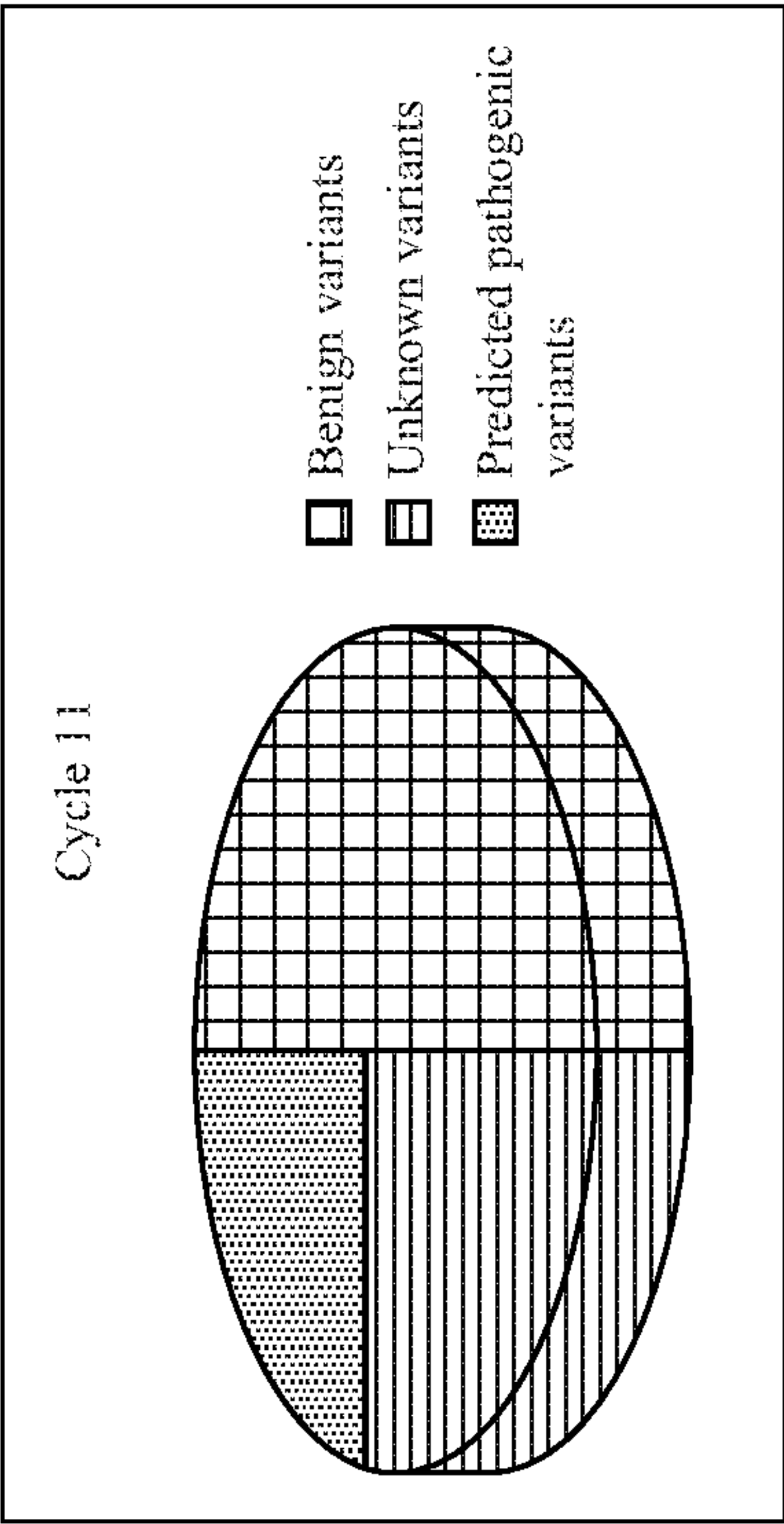
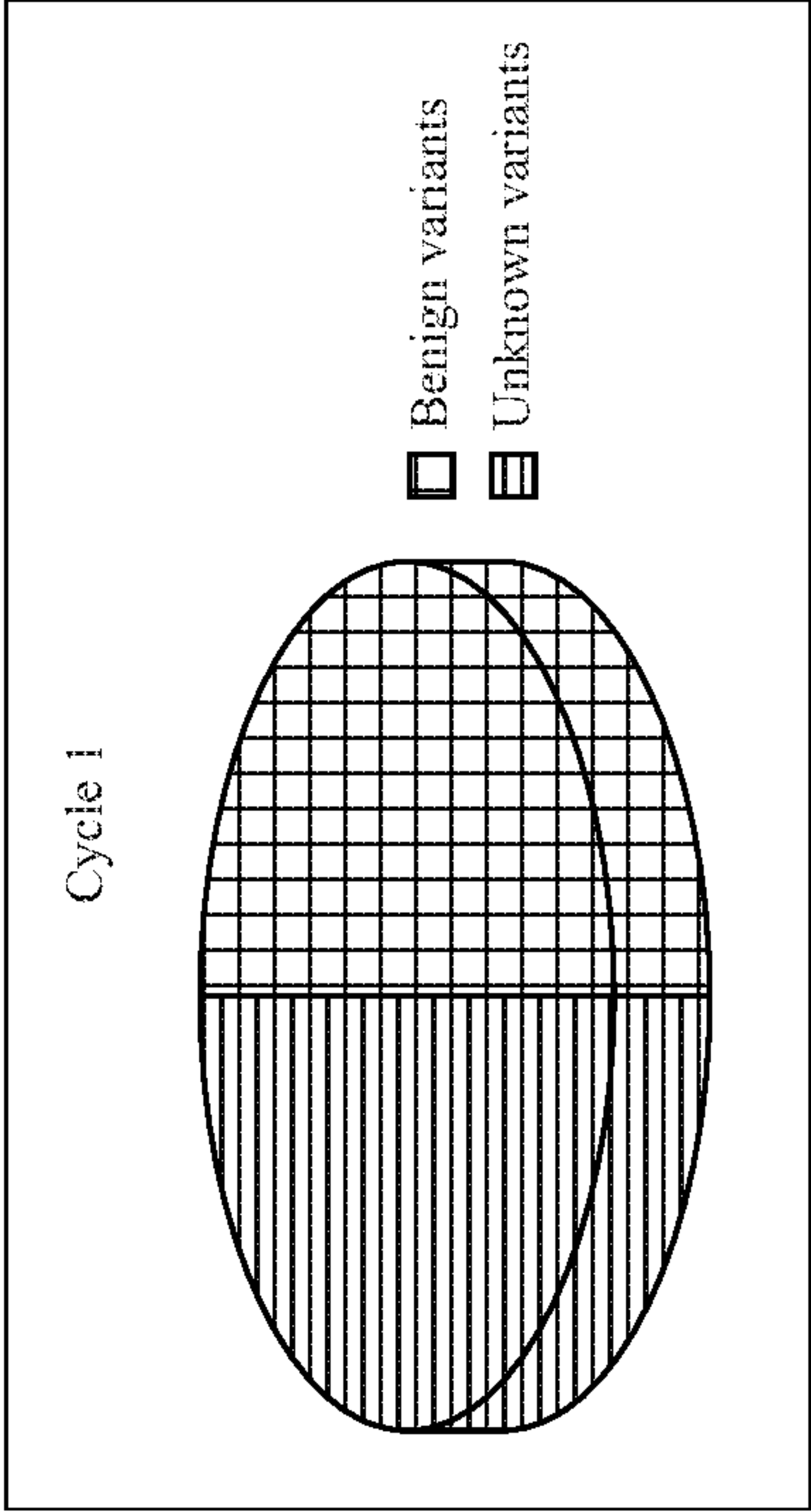


FIG. 42



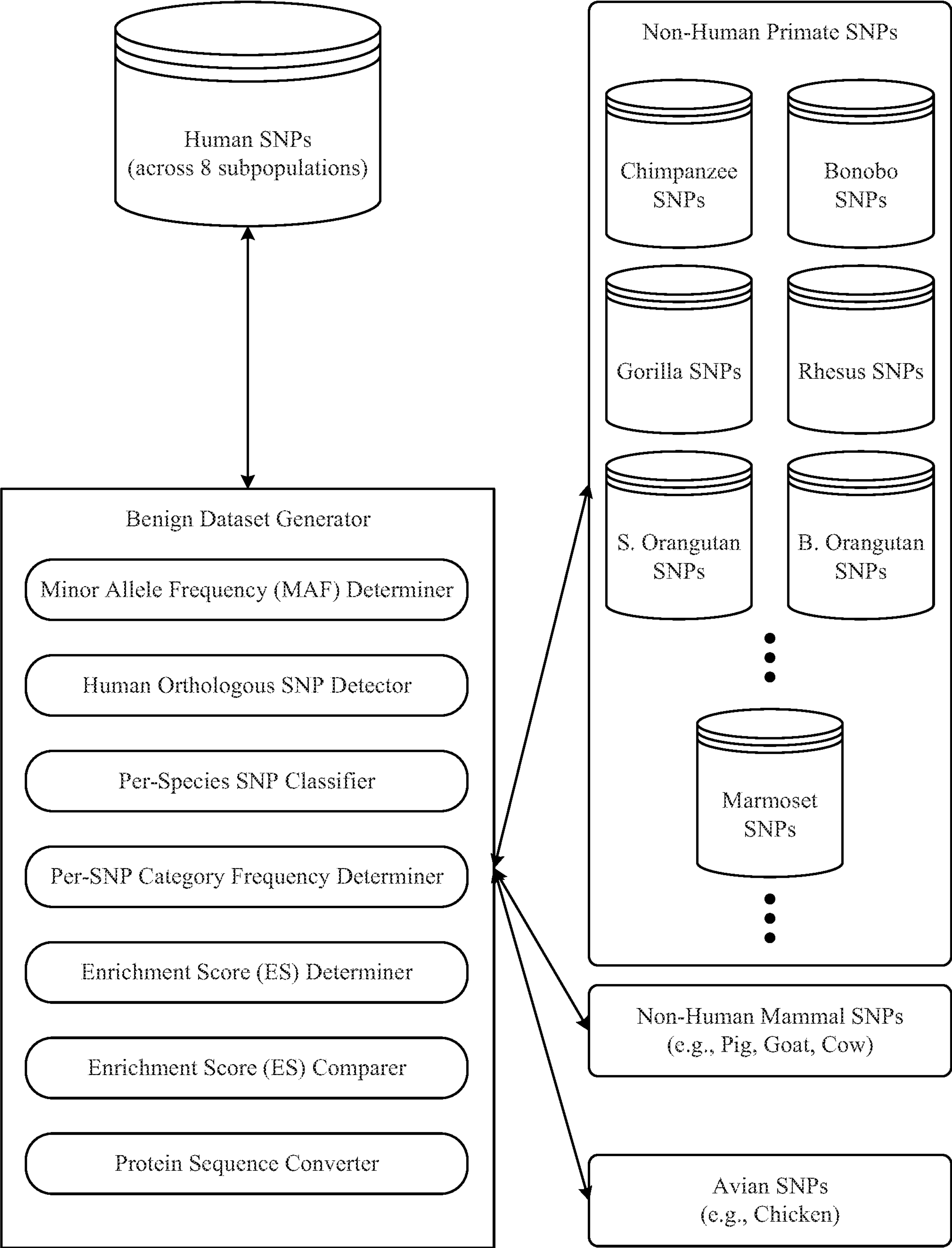


FIG. 43

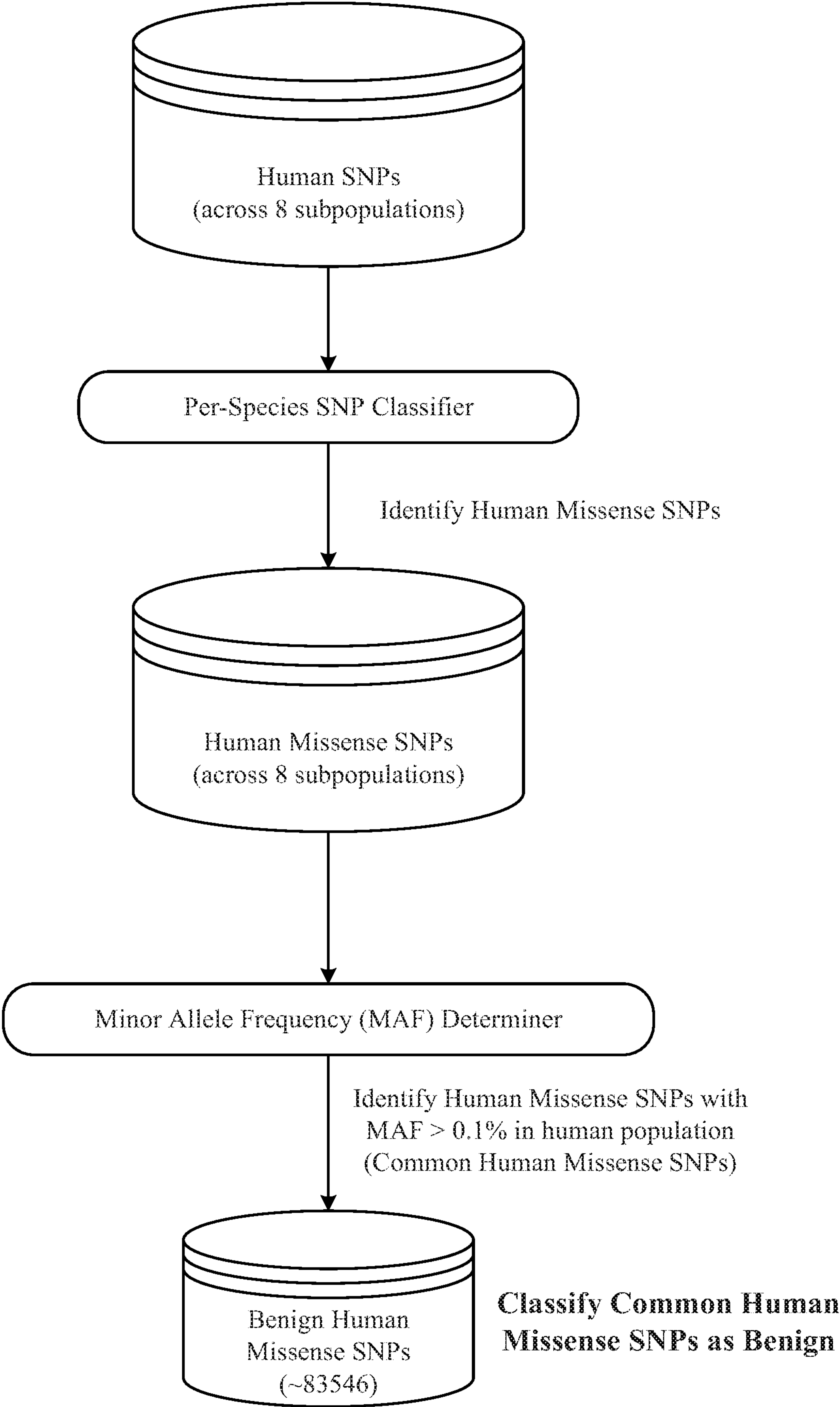
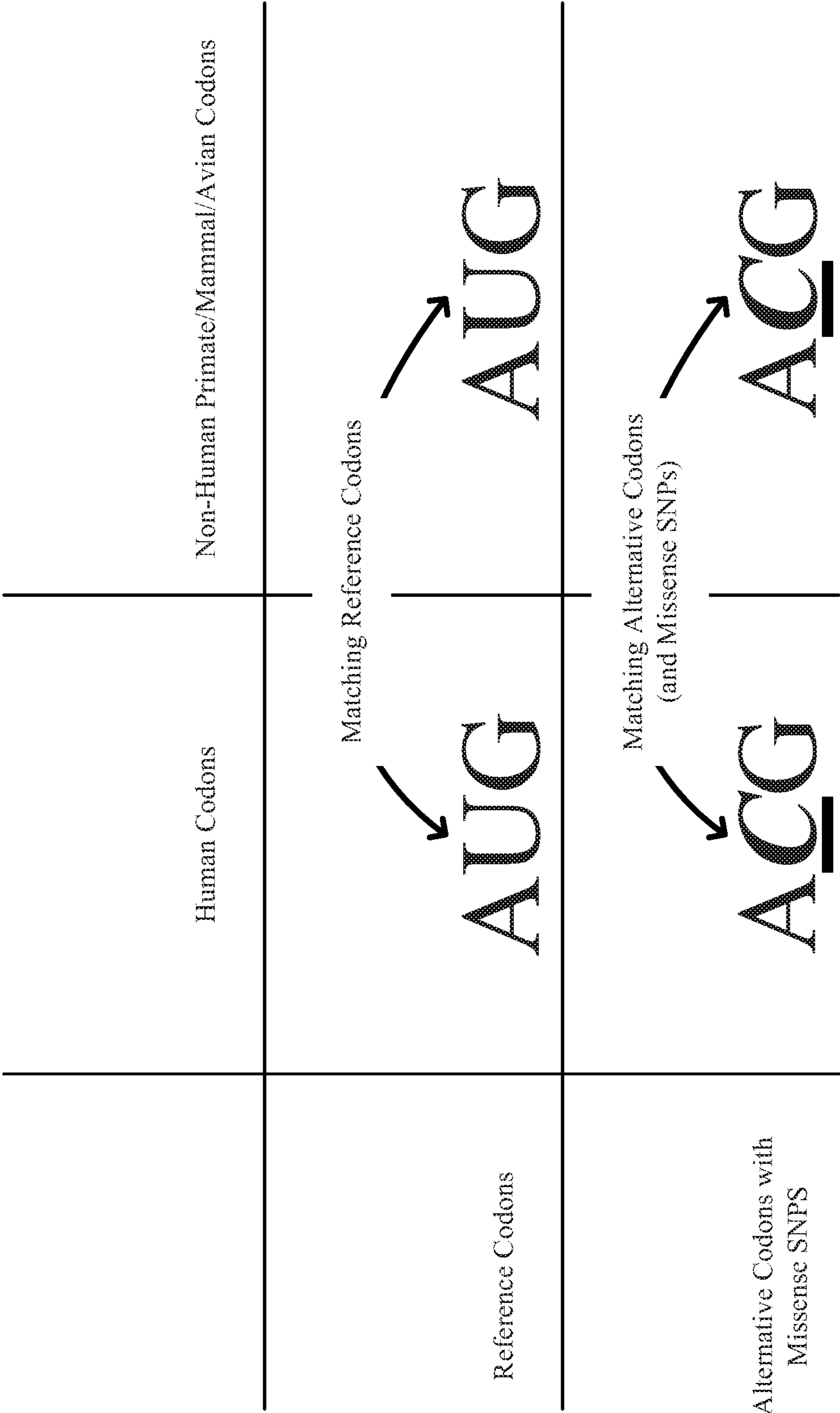


FIG. 44



Human Orthologous Missense SNP: A missense SNP in a non-human species that has matching reference and alternative codons with humans.

FIG. 45

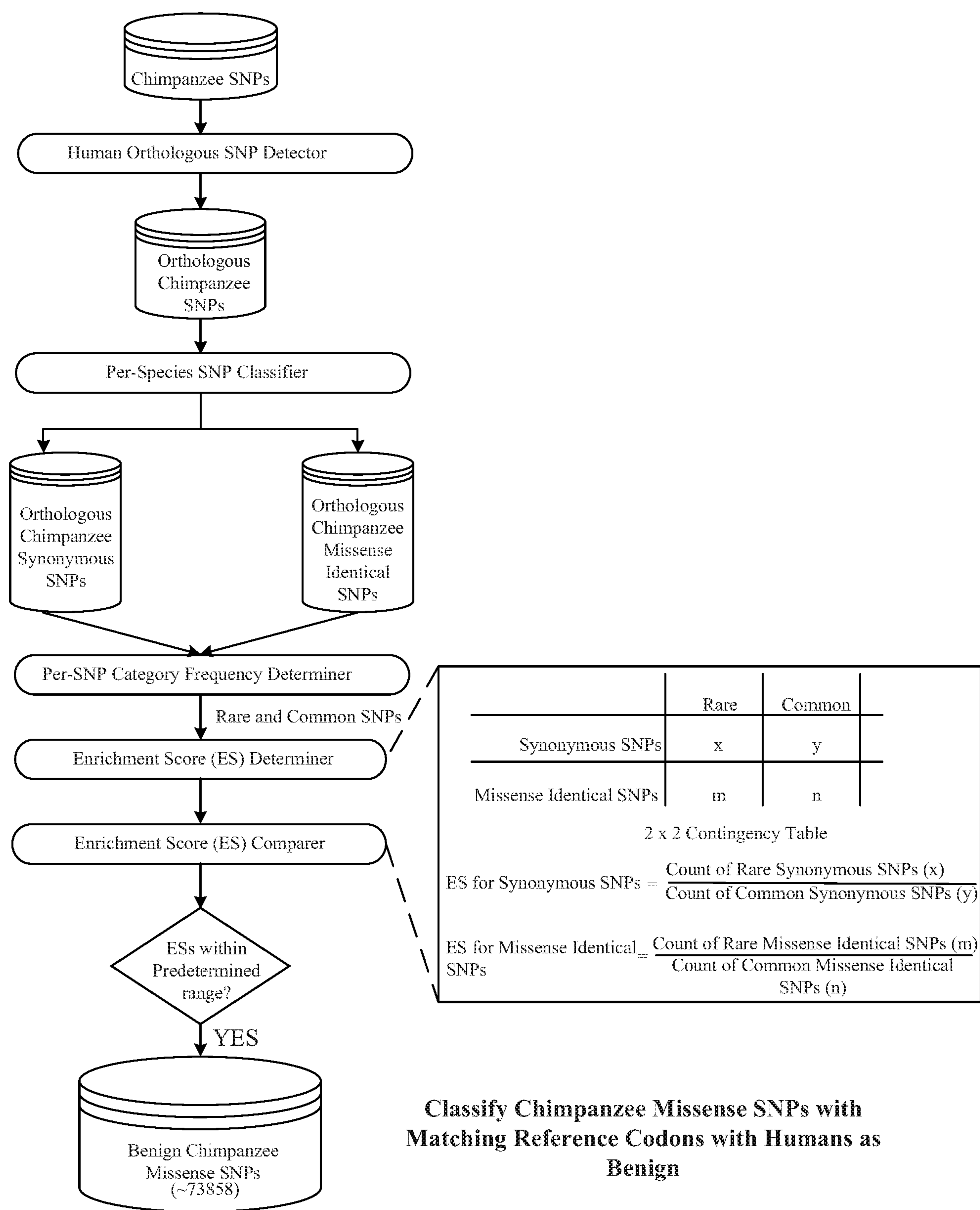


FIG. 46

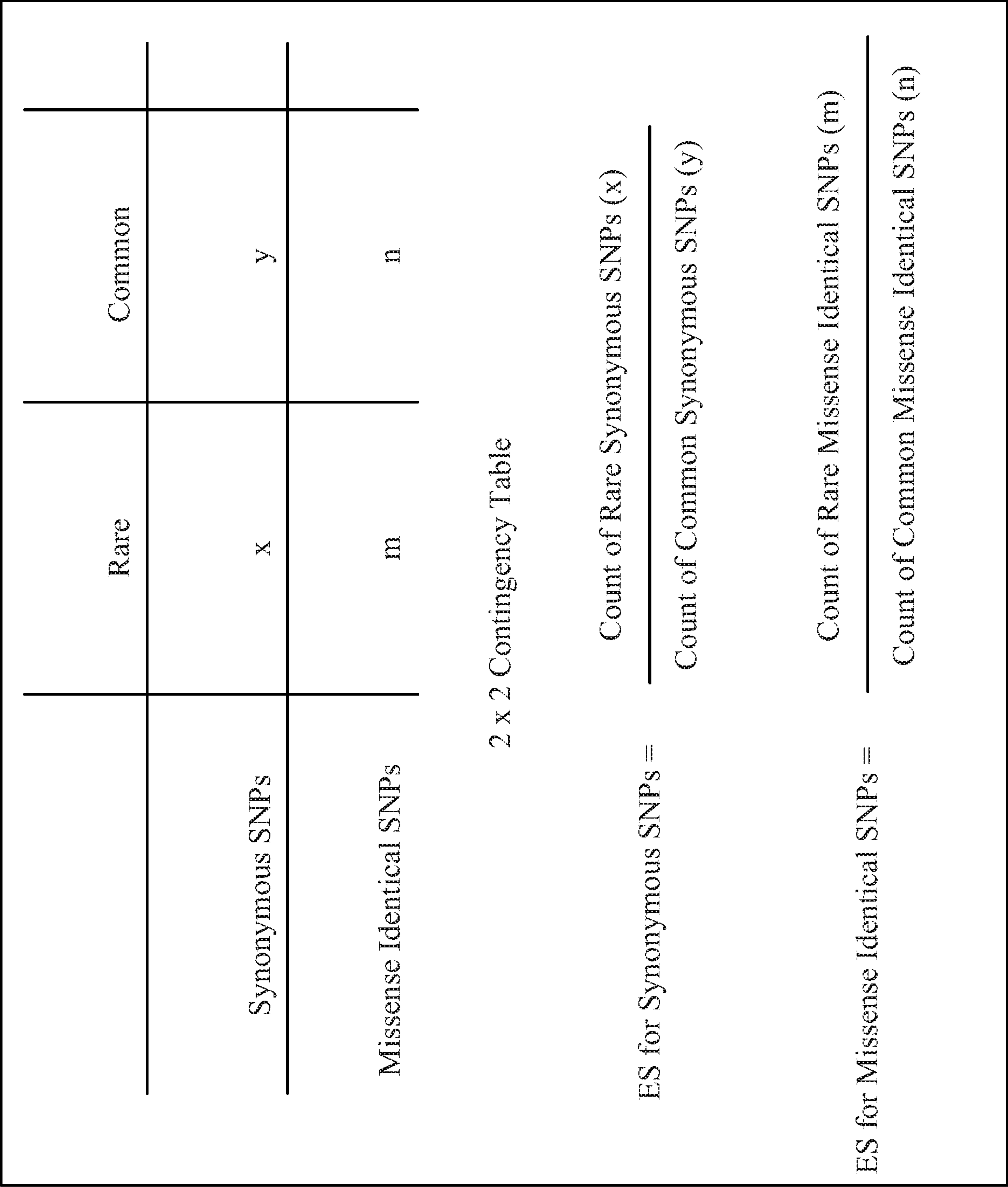


FIG. 47

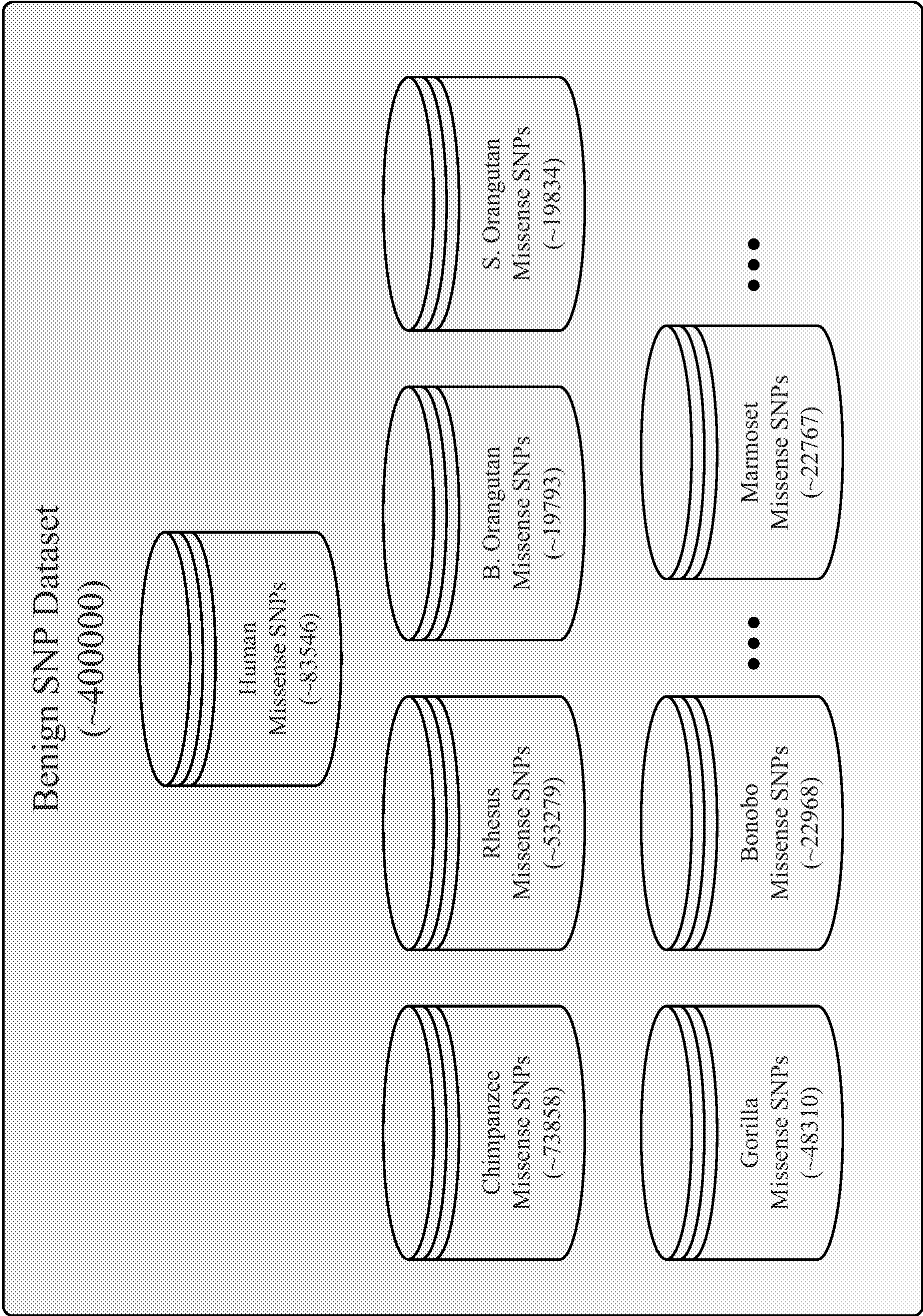


FIG. 48

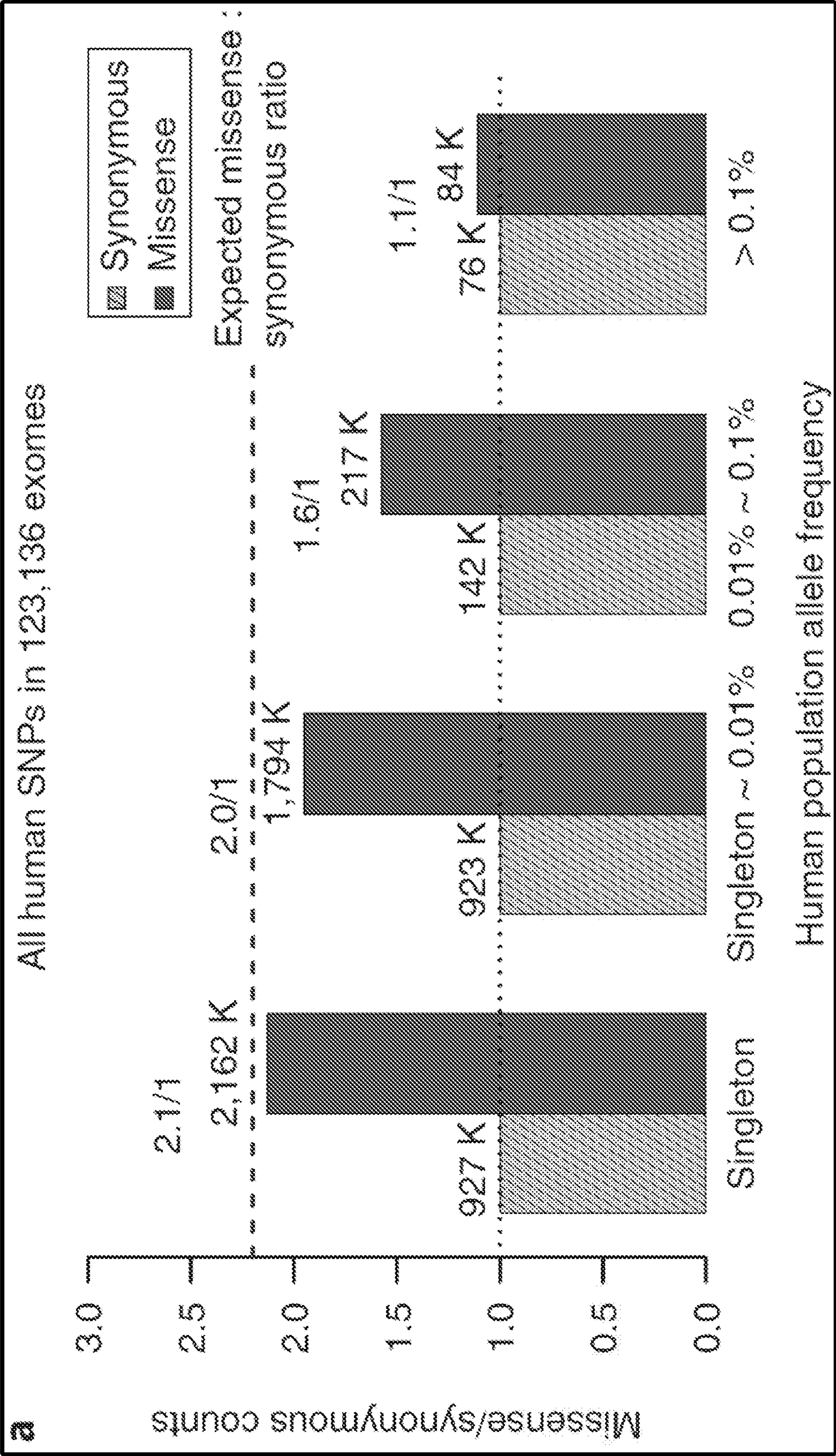


FIG. 49A

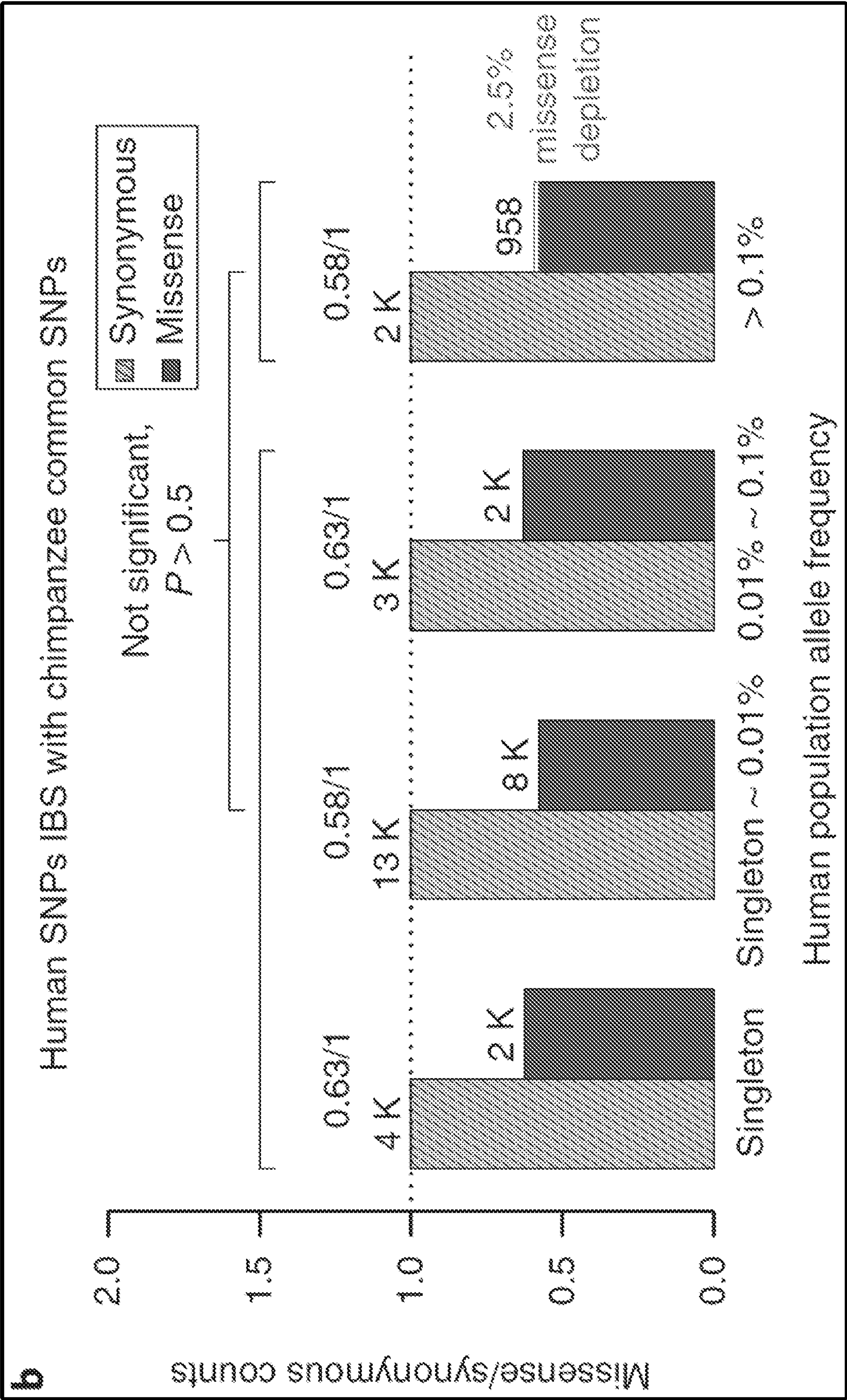


FIG. 49B

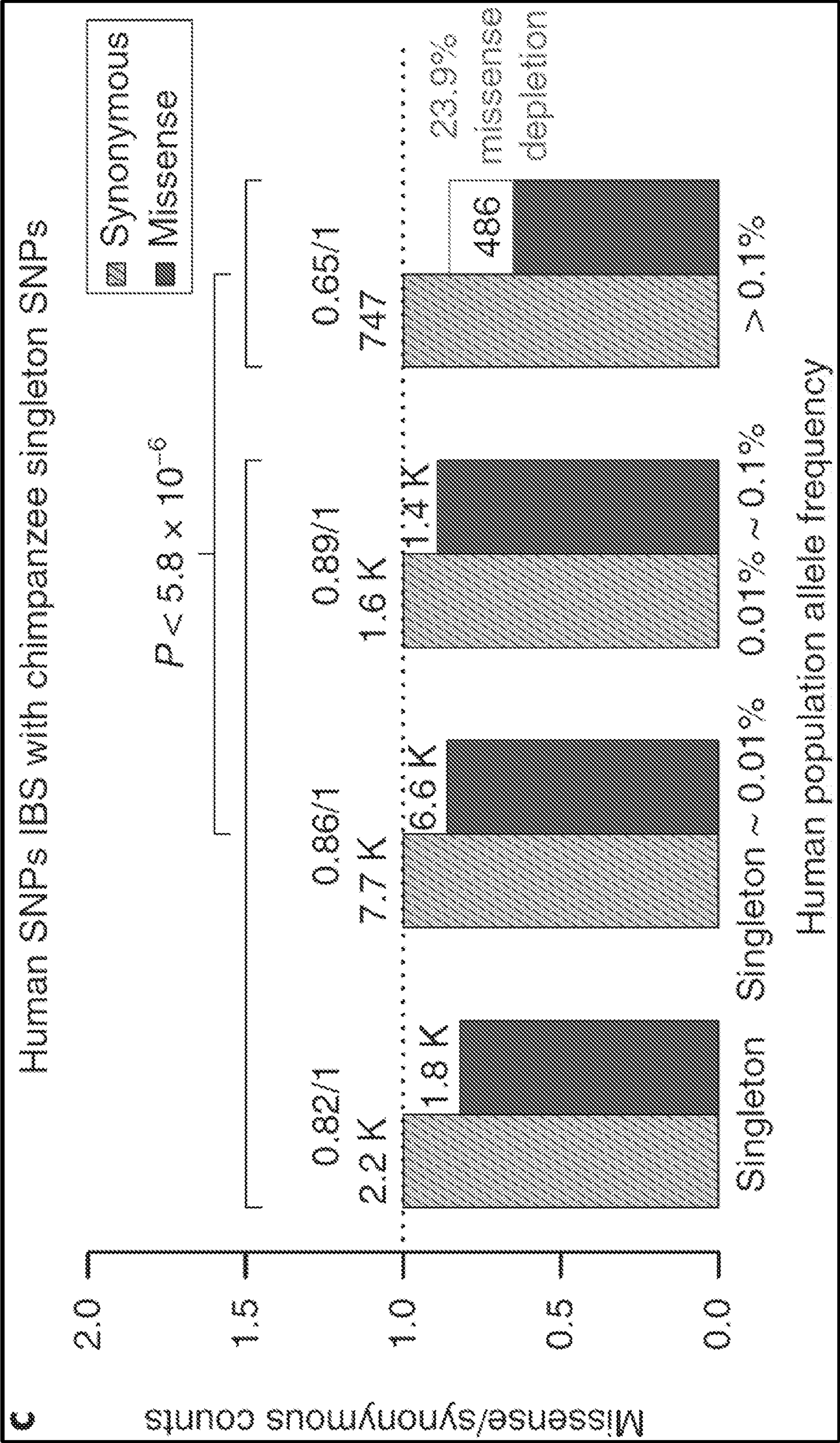


FIG. 49C

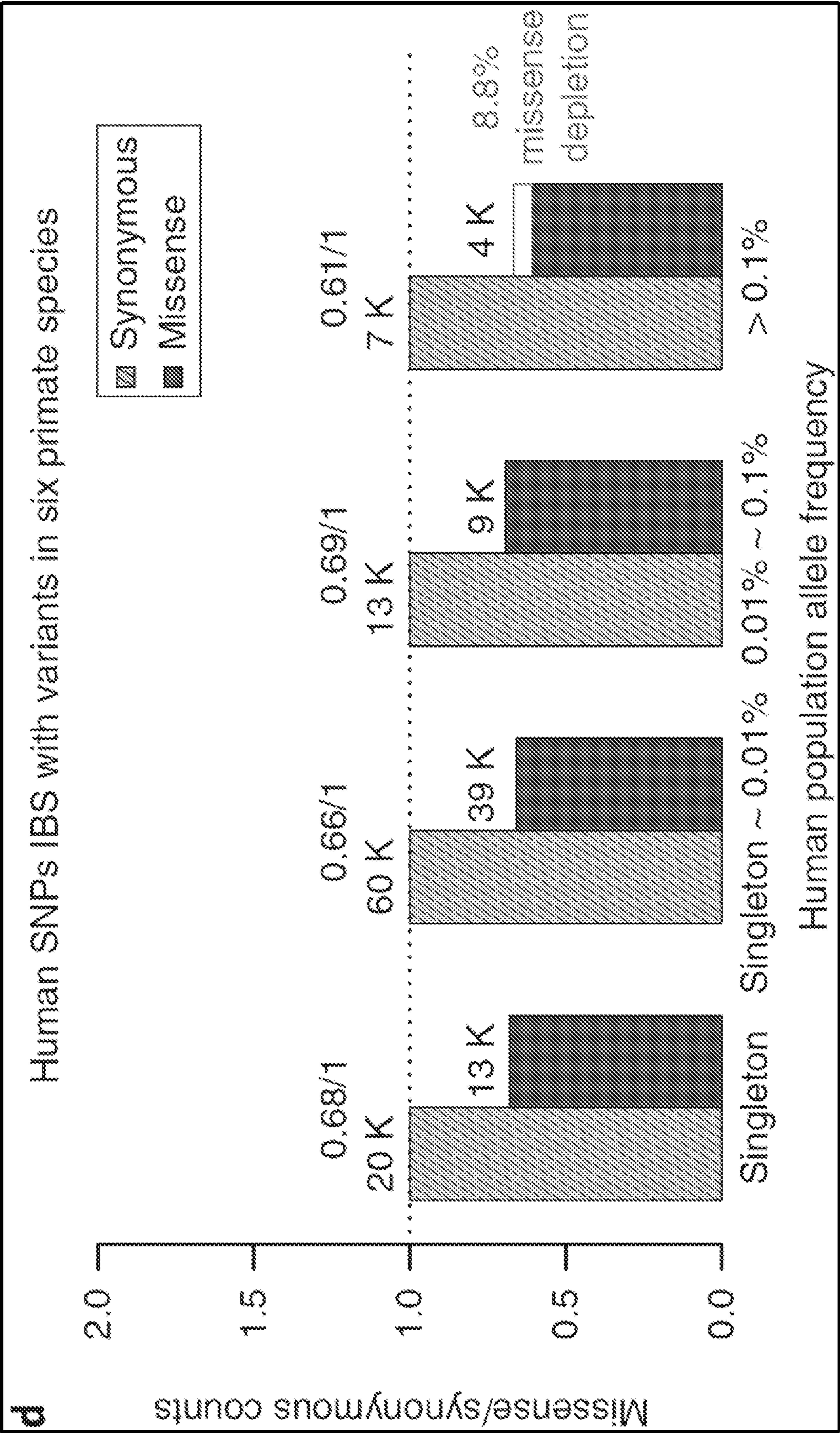


FIG. 49D

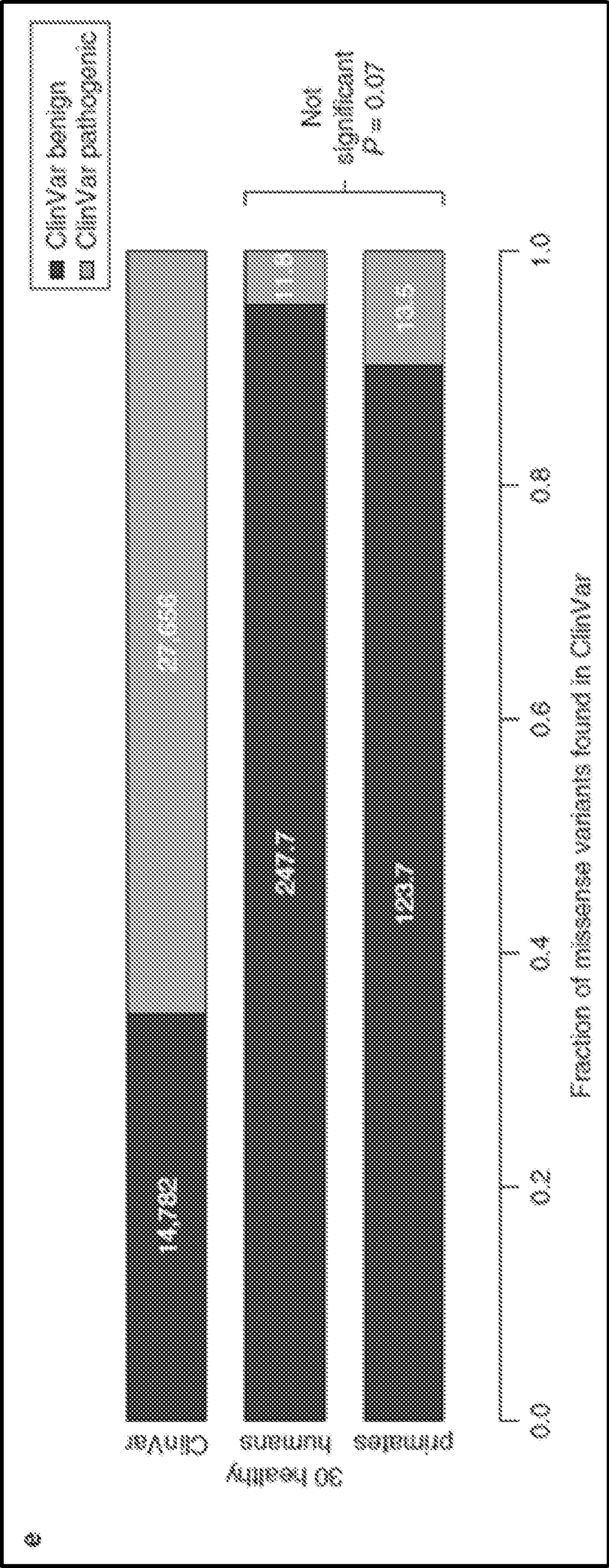
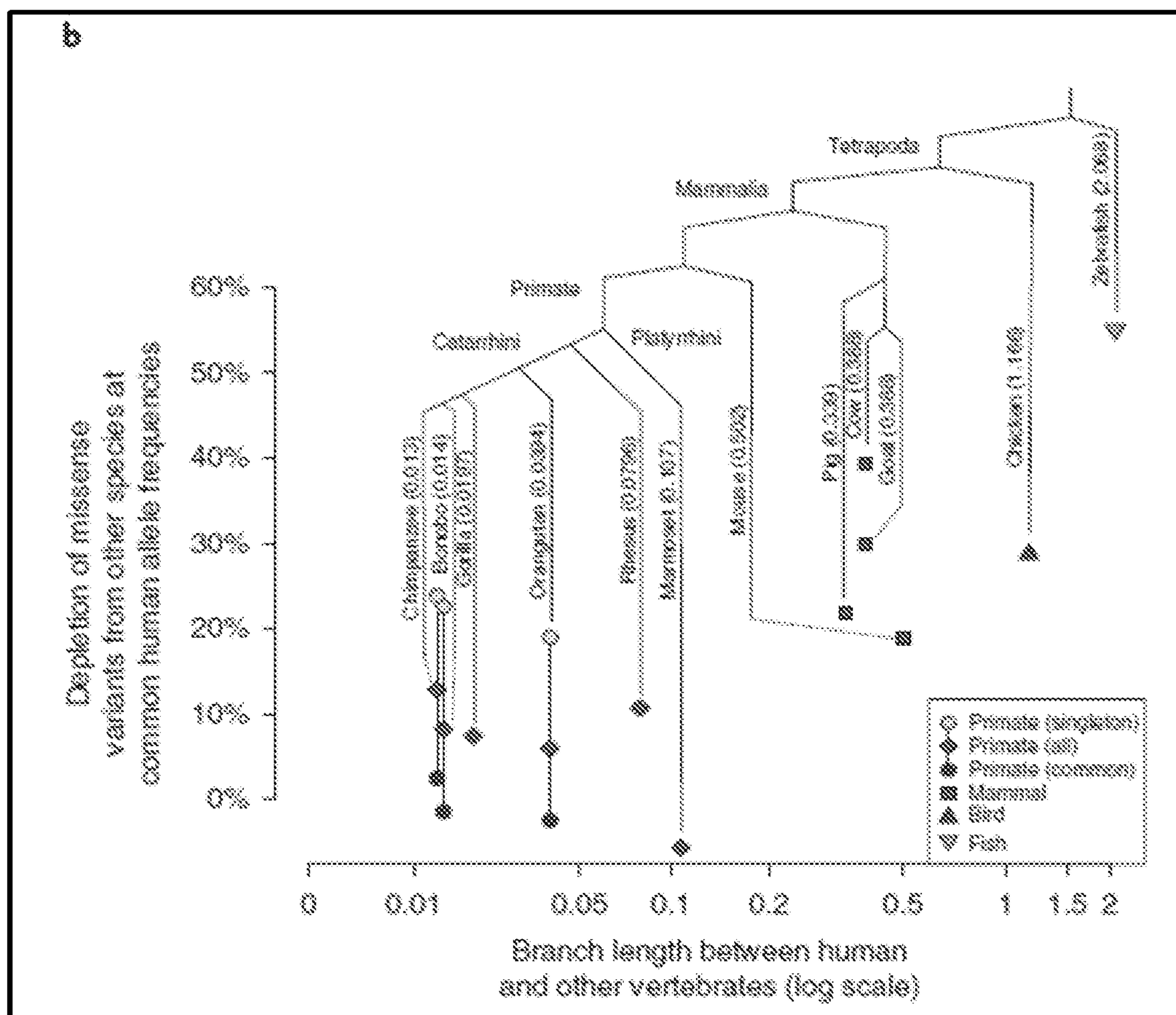
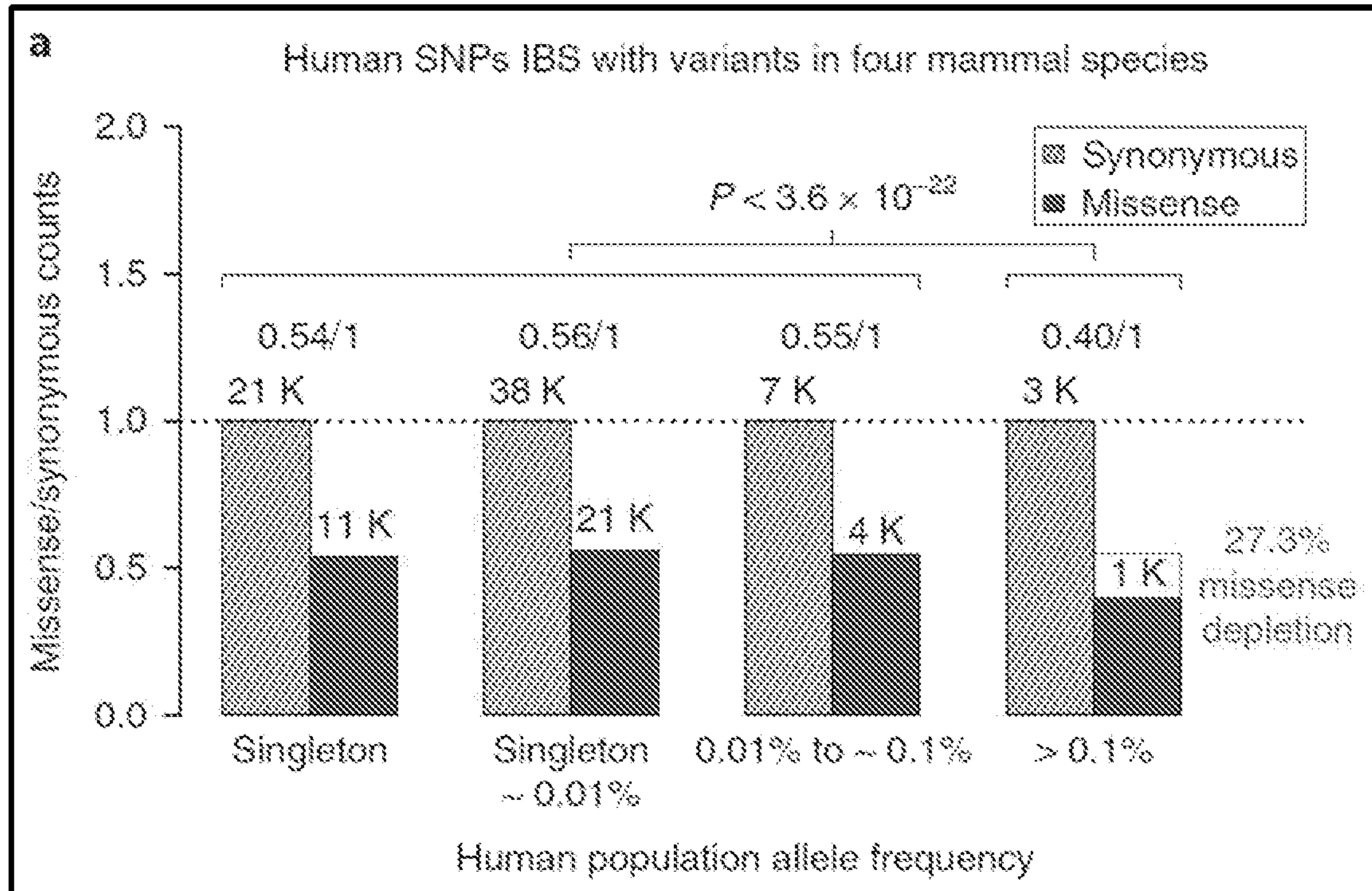
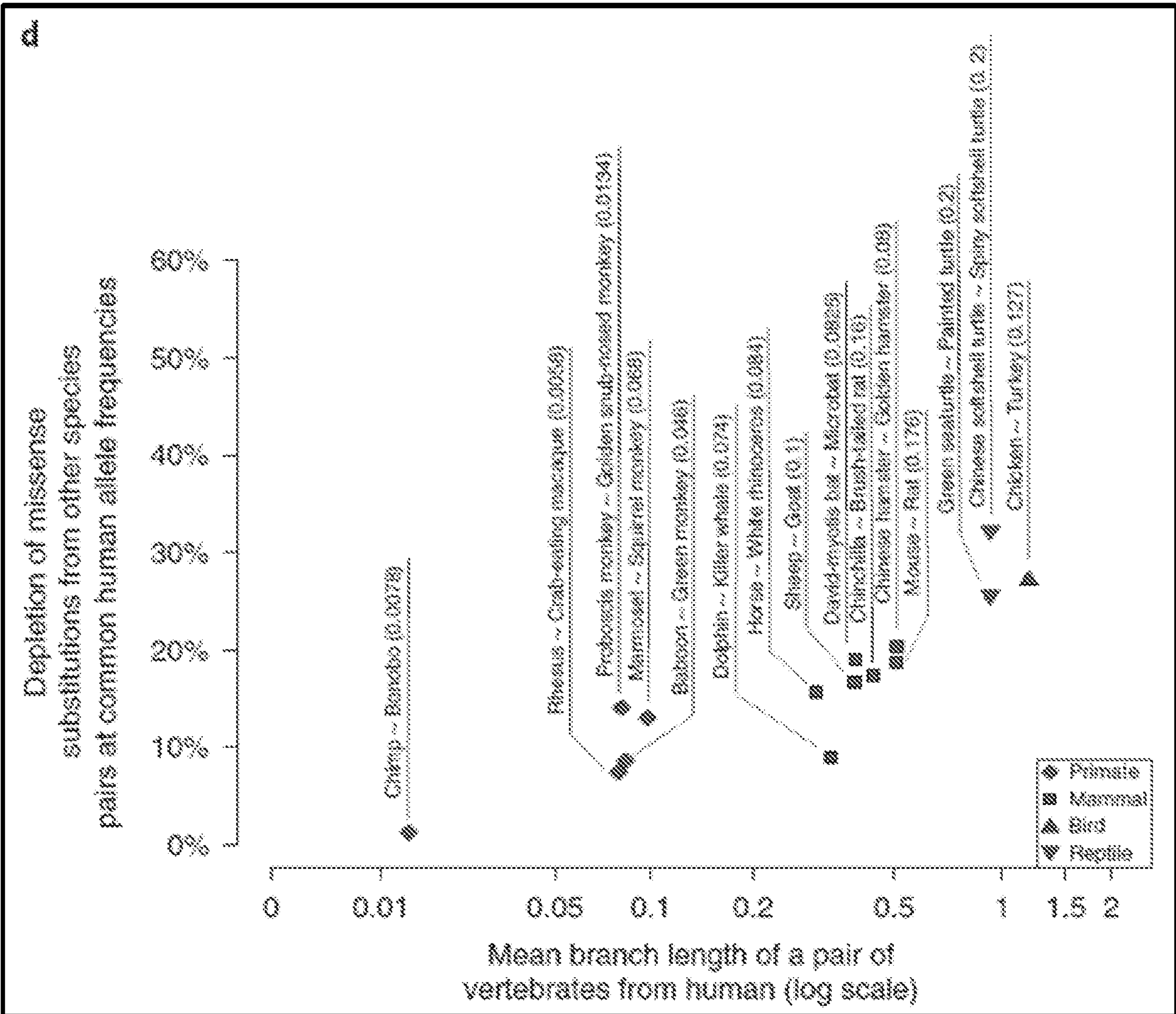
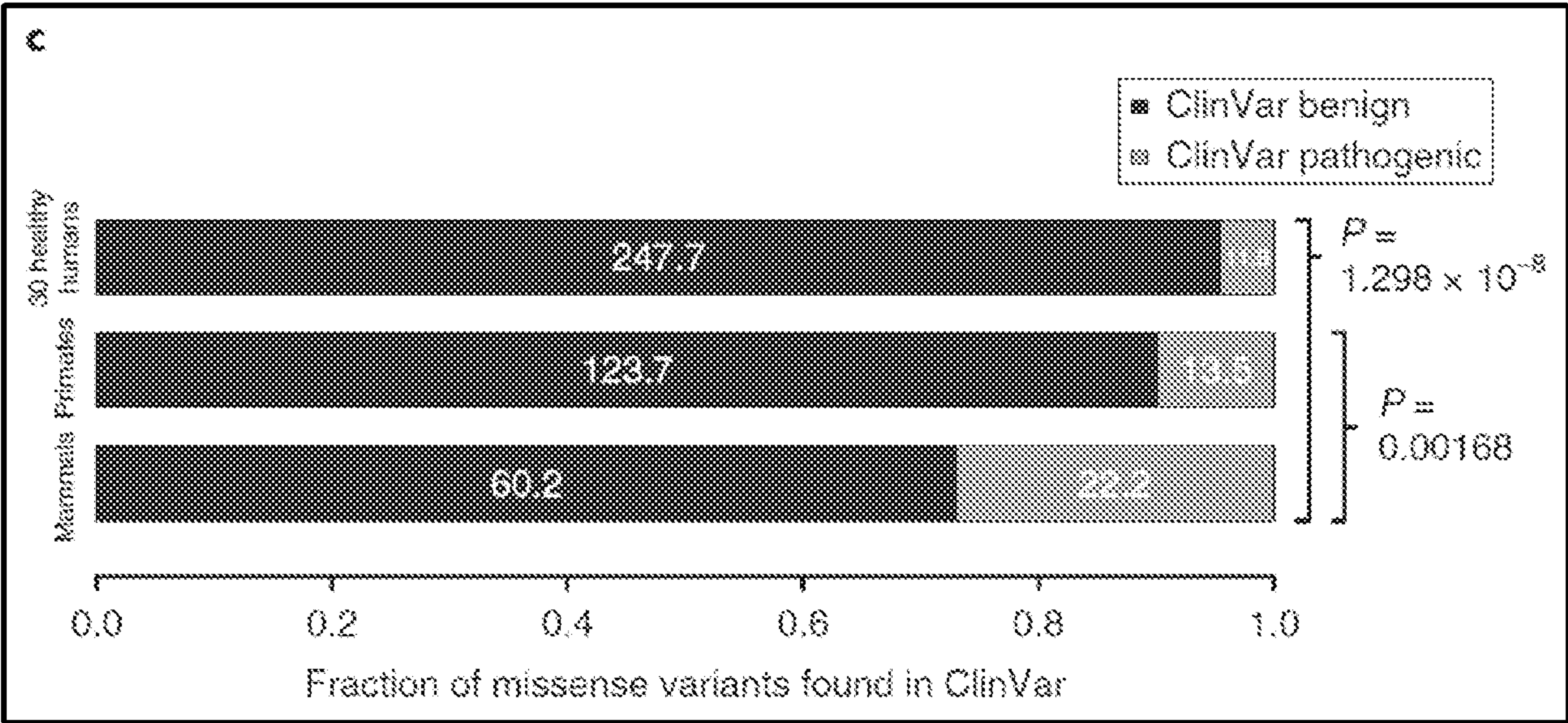


FIG. 49E



FIGs. 50A, 50B



FIGs. 50C, 50D

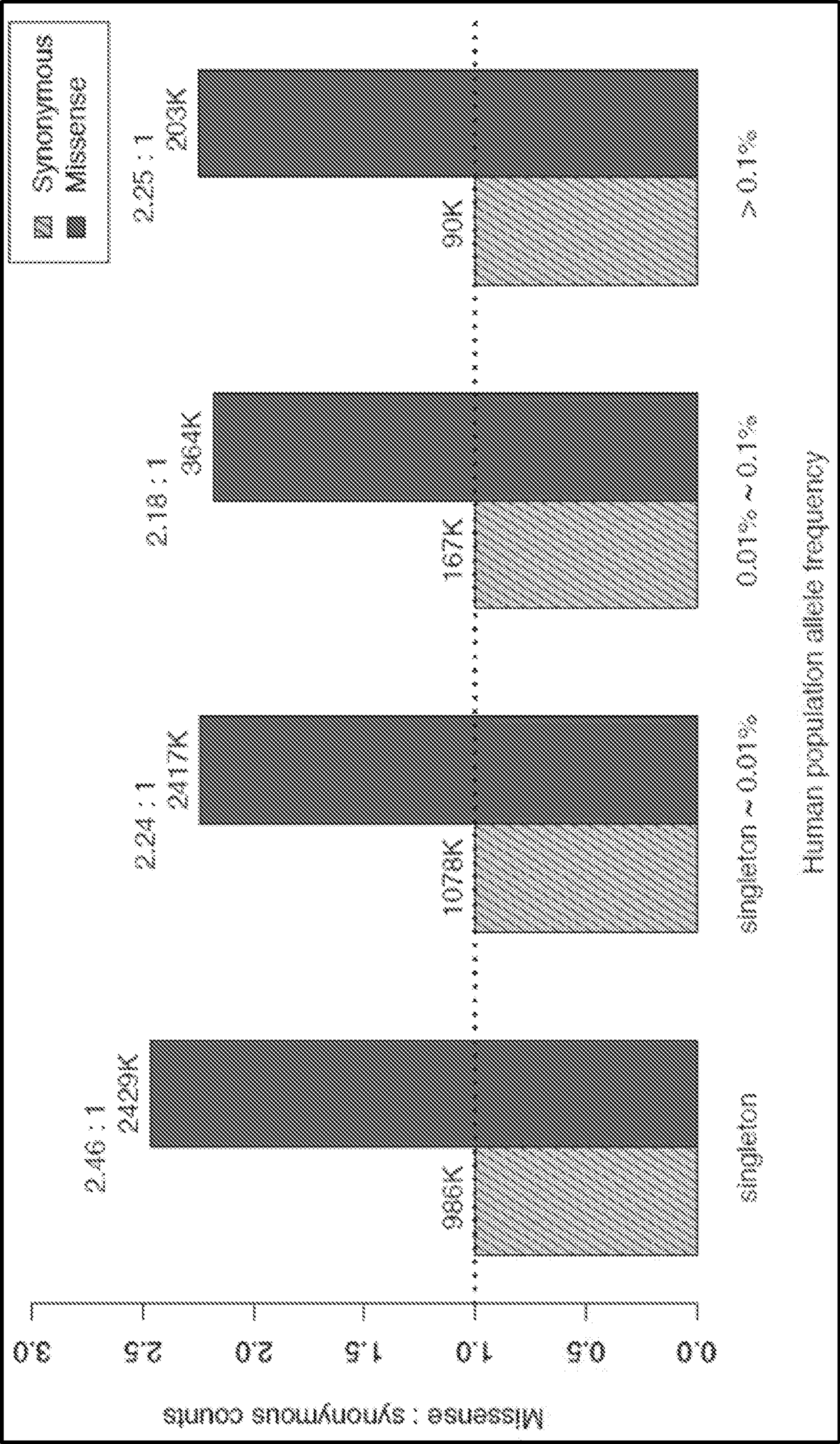


FIG. 51

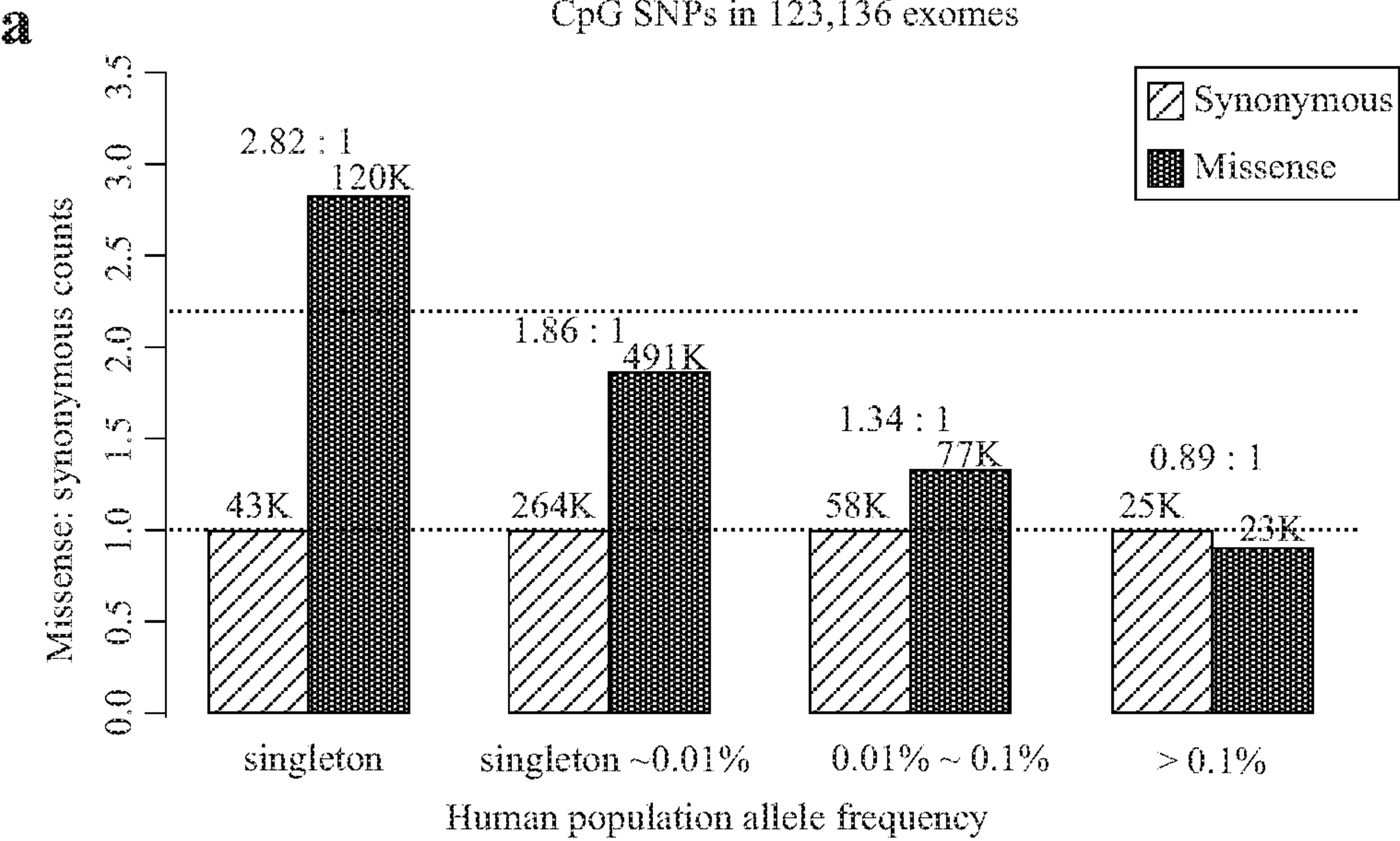


FIG. 52A

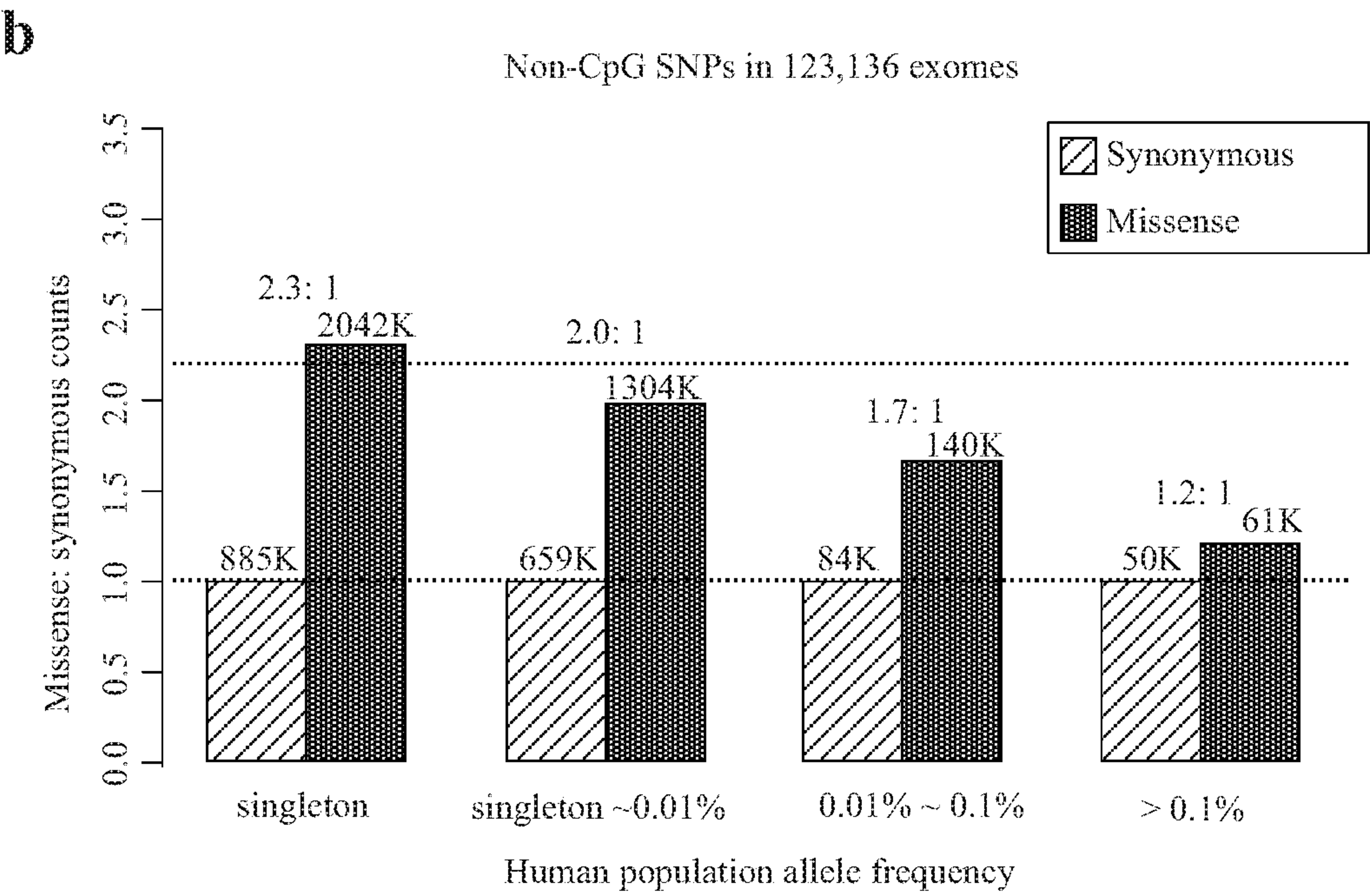


FIG. 52B

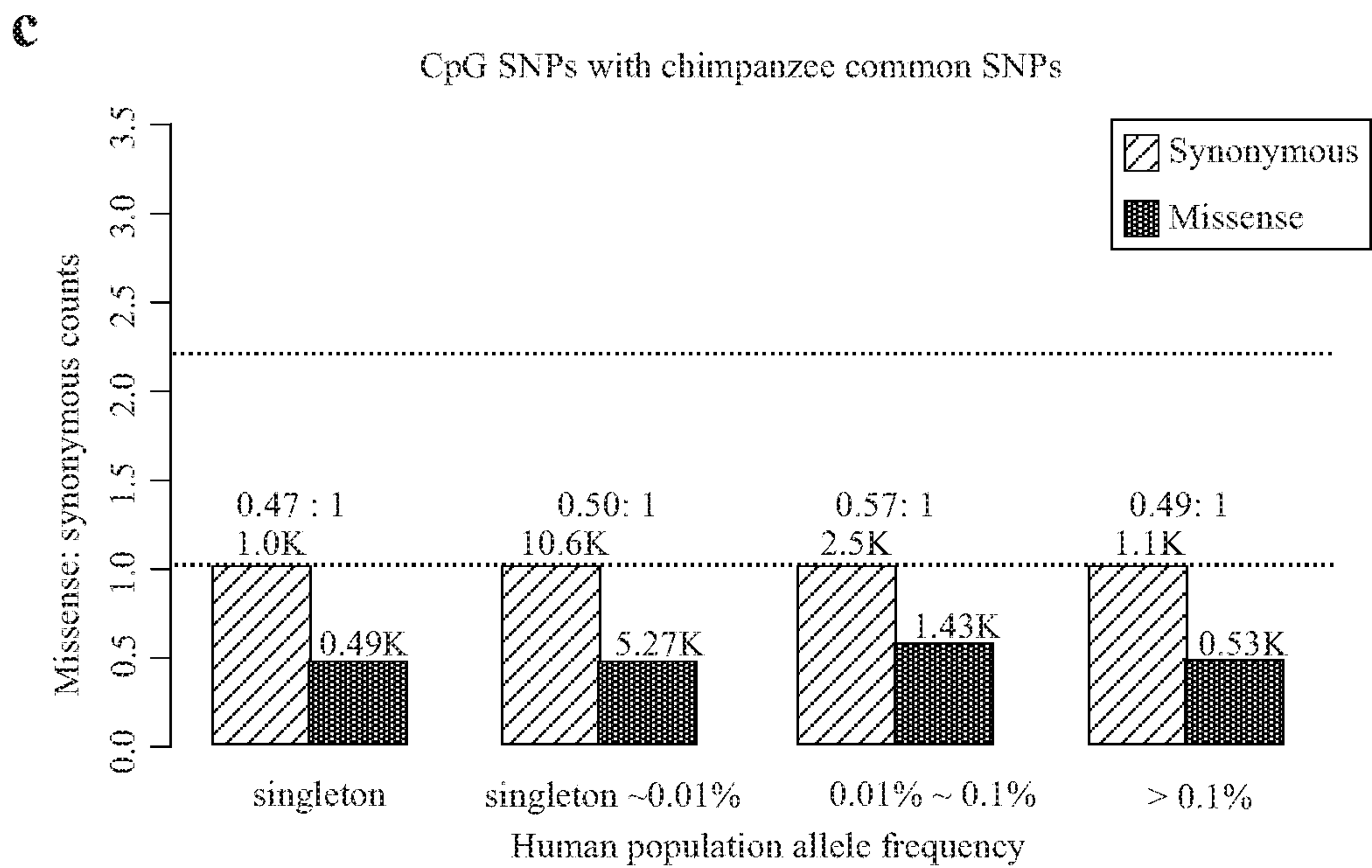


FIG. 52C

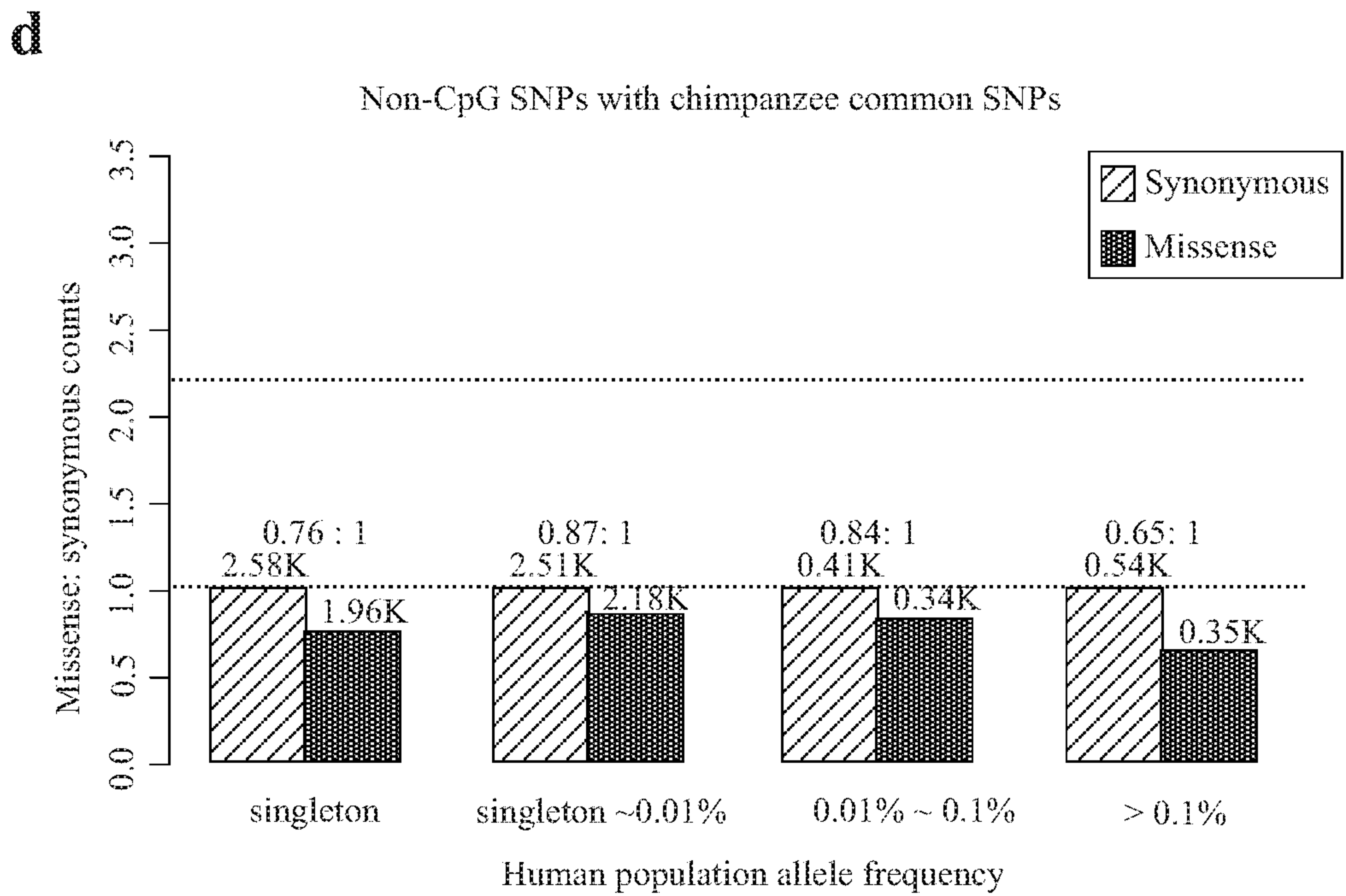


FIG. 52D

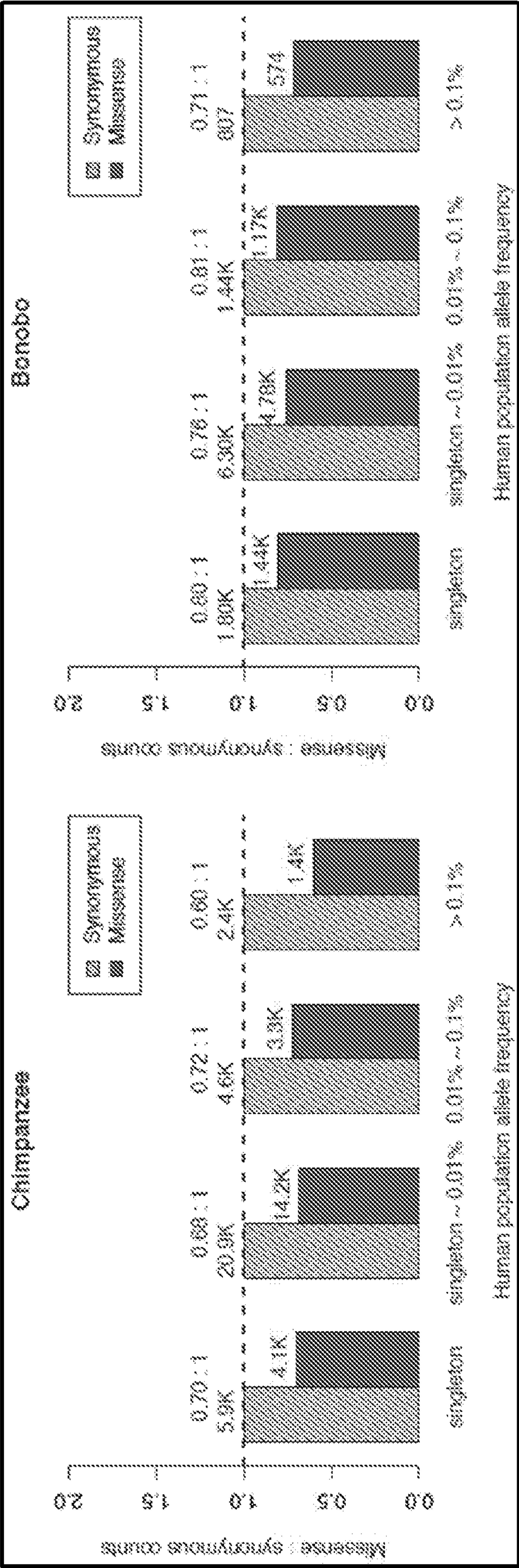


FIG. 53

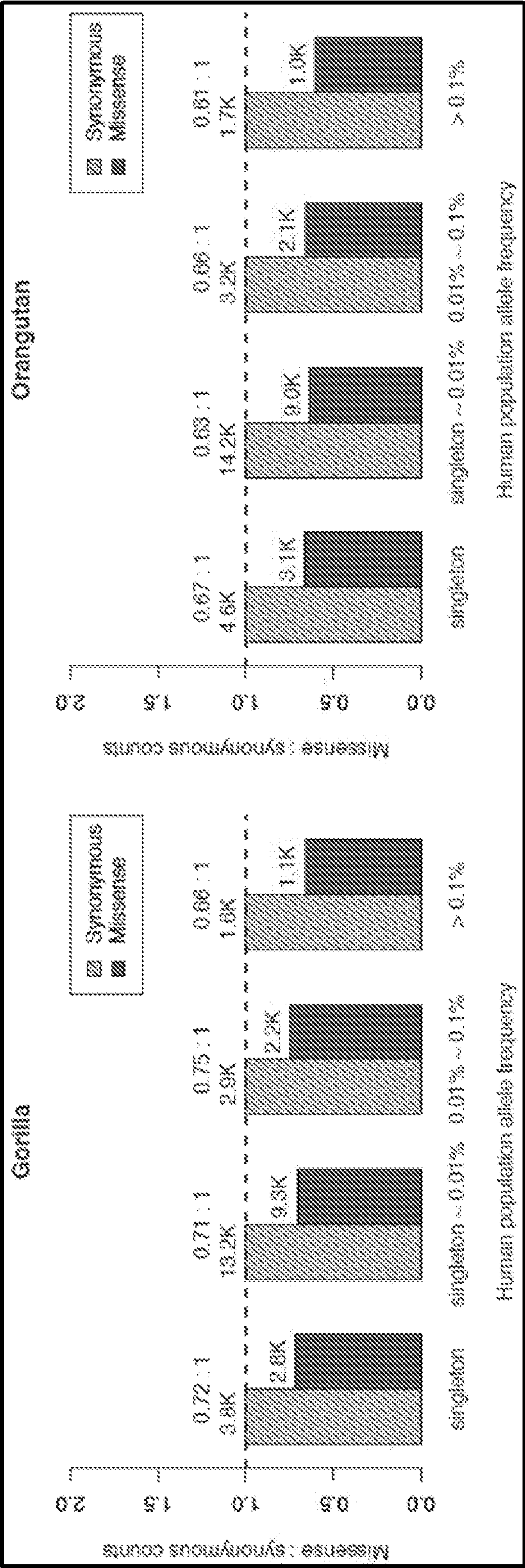


FIG. 54

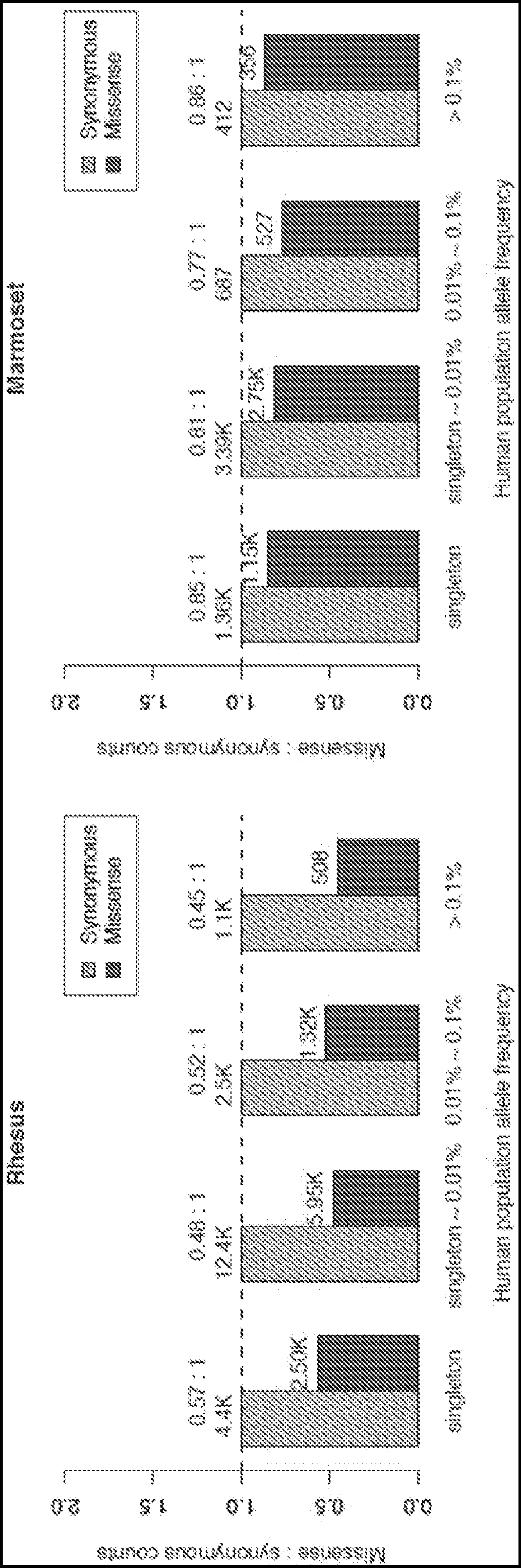


FIG. 55

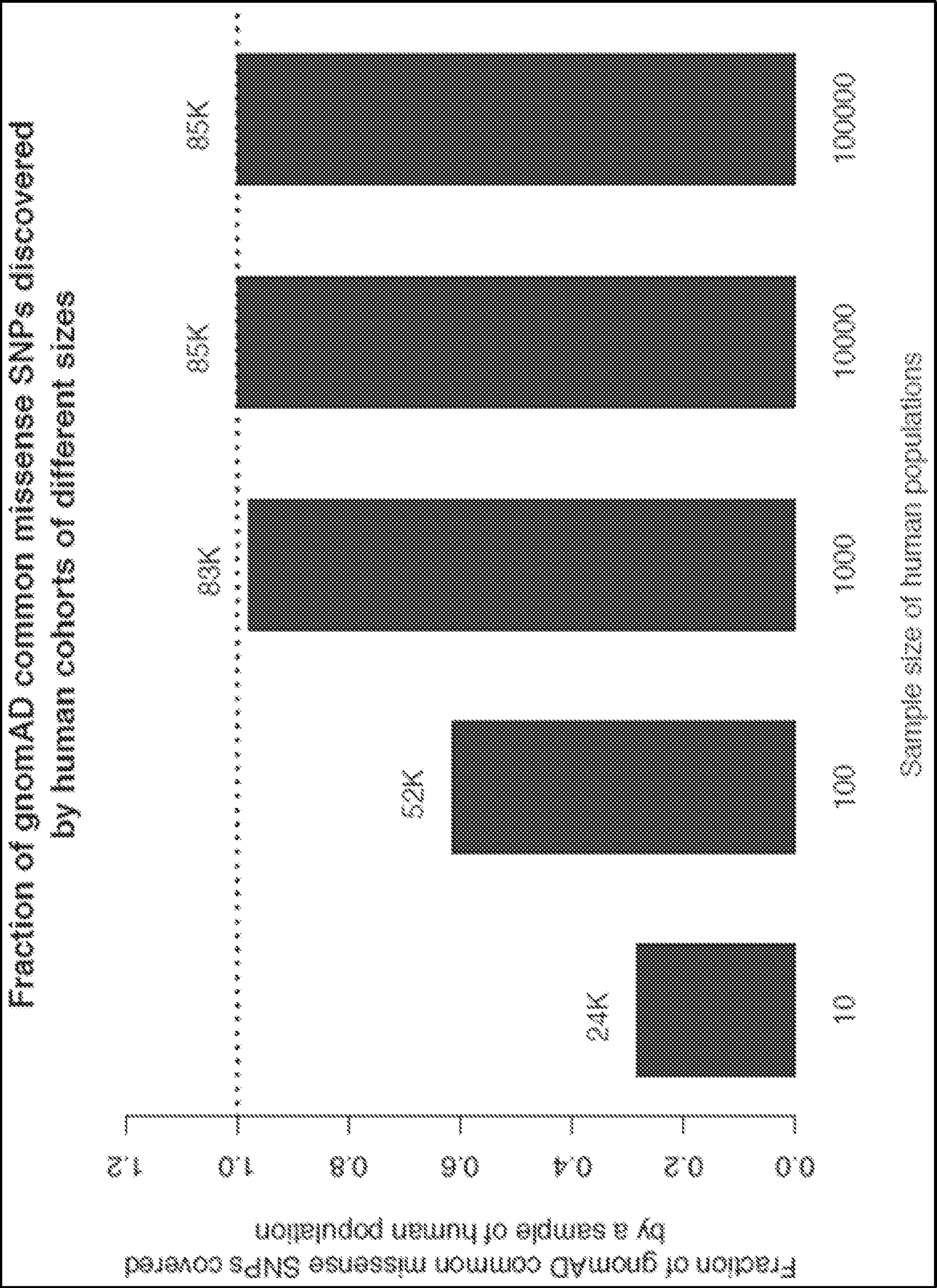


FIG. 56

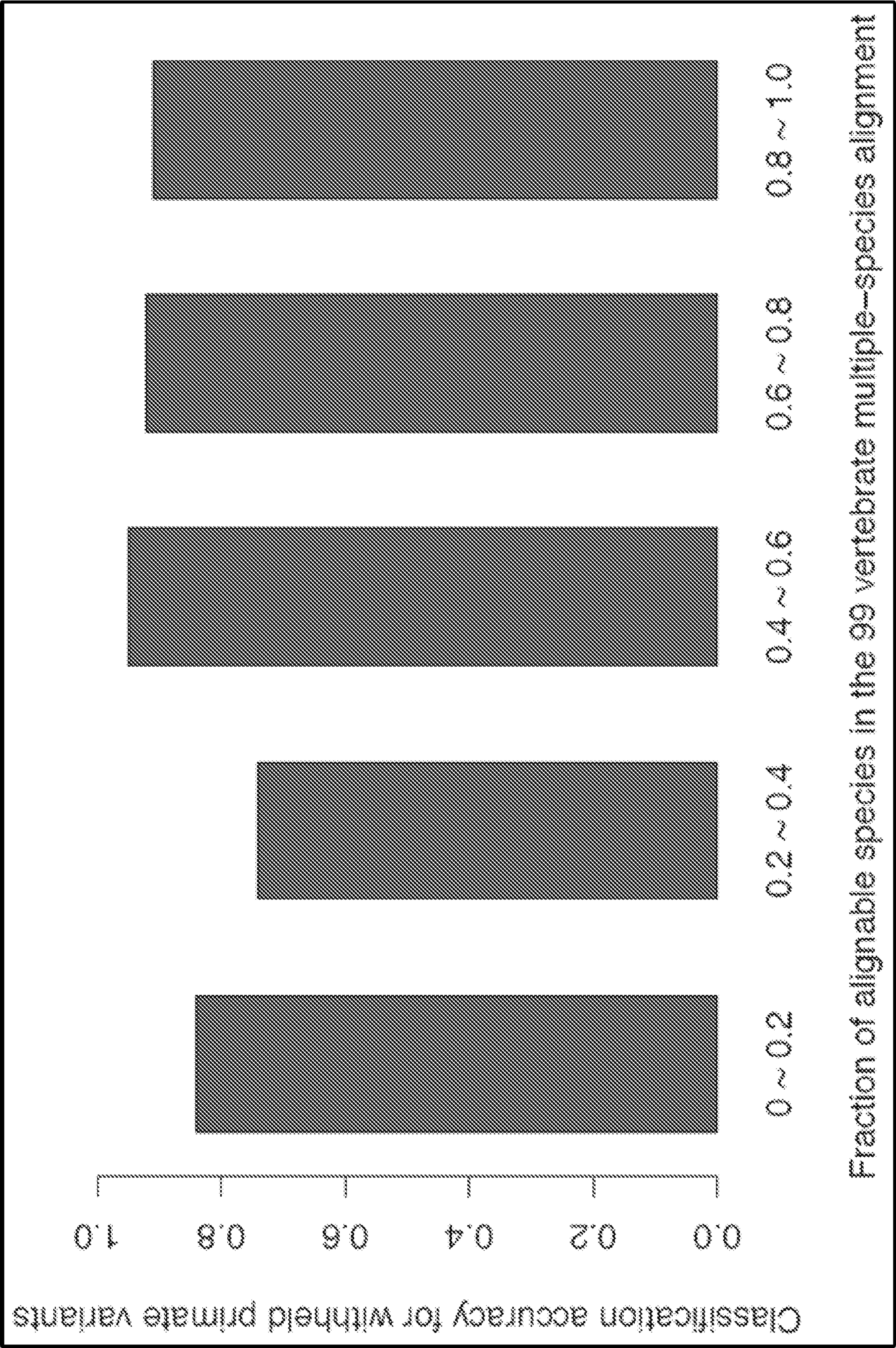


FIG. 57

Species (and number of individuals)	Source	Number of missense variants matching human codons	Number of unique missense variants contributed
Human common variants (MAF > 0.1%, 123136 individuals)	gnomAD / ExAC database	83,546	83,546
Chimpanzee (24)	Prado-Martinez <i>et al.</i> , Nature (2013)	76,293	73,858
Bonobo (13)	Prado-Martinez <i>et al.</i> , Nature (2013)	25,277	22,968
Gorilla (27)	Prado-Martinez <i>et al.</i> , Nature (2013)	52,052	48,310
Orangutan, Bornean subspecies (5)	Prado-Martinez <i>et al.</i> , Nature (2013)	21,621	19,793
Orangutan, Sumatran subspecies (5)	Prado-Martinez <i>et al.</i> , Nature (2013)	27,567	19,834
Chimpanzee (35)	de Manuel <i>et al.</i> , Science (2016)	75,580	30,327
Orangutan (individual data not available)	dbSNP	18,627	10,554
Rhesus (16, individual data not available)	dbSNP	62,393	53,279
Marmoset (9, individual data not available)	dbSNP	25,048	22,767
Total			385,236

FIG. 58

Human population allele frequency in 123,136 exomes	Expected number of synonymous variants	Expected number of missense variants	Missense: synonymous ratio
singleton	986,141.22	2,429,287.85	2.46
singletons~0.01%	1,077,630.37	2,416,751.61	2.24
0.01%~0.1%	166,849.26	364,292.75	2.18
>0.1%	90,221.42	202,918.38	2.24

FIG. 59

	Total number of ClinVar variants of interest
Original ClinVar file	666,403
Retaining only hg19 records	324,698
Extracting only single nucleotide variants	264,623
Removing non-coding variants	186,769
Ignoring all synonymous changes and missense changes creating or disrupting a stop codon	122,884
Excluding variants of unknown significance and variants with conflicting annotations	42,438

FIG. 60

Species	Number of variants with benign annotations in ClinVar	Number of variants with pathogenic annotations in ClinVar
Chimpanzee	218	21
Bonobo	85	7
Gorilla	167	23
Orangutan	160	12
Rhesus	87	11
Marmoset	25	7
Mouse	30	4
Pig	74	30
Cow	77	39
Goat	60	16
Chicken	9	8
Zebrafish	2	6

FIG. 61

Table 1 | Additional genes achieving genome-wide significance in intellectual disability when considering only missense de novo mutations (DNMs) with PrimateAI scores ≥ 0.803

HGNC symbol	Protein-truncating variants	Missense	P value	Phenotypic abnormalities observed in multiple individuals		
		PrimateAI score ≥ 0.803	All missense	PrimateAI score ≥ 0.803	All missense	
ACTL68	0	3	1.5×10^{-7}	3	2.4×10^{-6}	Microcephaly
EBF3	3	3	5.2×10^{-8}	3	5.4×10^{-6}	Growth delay, eye abnormality, strabismus, ataxia
EFTUD2	2	4	1.5×10^{-7}	4	1.5×10^{-5}	Microcephaly, low-set ears, microtia, choanal atresia
HECW2	1	8	2.8×10^{-10}	8	6.7×10^{-7}	Seizures, myopathy, abnormal calvarium
KDM6A	2	3	2.3×10^{-7}	3	9.8×10^{-6}	Eyelid, dental abnormalities, hypotonia
KIF5C	0	3	3.0×10^{-7}	3	2.8×10^{-6}	Cerebral hypoplasia
MAP2K1	0	5	3.1×10^{-8}	5	2.7×10^{-6}	Hypertelorism, low-set ears, polyhydramnios
PPP1CB	0	6	1.5×10^{-8}	6	1.6×10^{-6}	Abnormality of the forehead, short stature
PRKD1	0	6	8.6×10^{-8}	6	1.7×10^{-5}	Skin, digital, and cardiac abnormalities; sparse hair
SOX11	1	3	3.1×10^{-7}	3	2.4×10^{-6}	Hypermetropia, nail hypoplasia
TBR1	4	4	1.3×10^{-6}	4	4.2×10^{-7}	Autistic behavior
TLK2	3	5	4.7×10^{-9}	5	6.3×10^{-7}	Nose, eyelid abnormalities, slanted palpebral fissure
TRIP12	6	2	1.4×10^{-7}	4	5.4×10^{-7}	Joint laxity
UZAF2	0	4	2.6×10^{-7}	4	1.2×10^{-6}	Seizures; eye, palatal, philtrum abnormalities

Counts of protein truncating and missense DNMs are provided. P values for gene enrichment are shown when the statistical test was run only with missense mutations with PrimateAI score ≥ 0.803 , and when it was repeated for all missense mutations.

FIG. 62

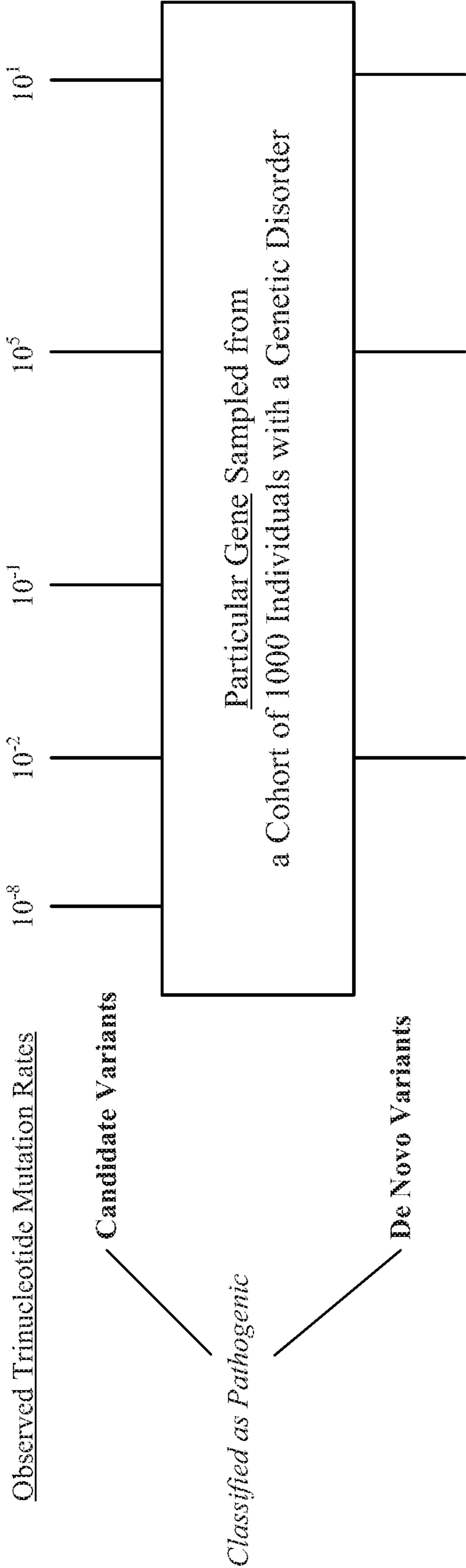
Table 2 | Comparison of the difference in Grantham score, protein surface-exposure, and amino acid sequence conservation between human expert annotated variants in ClinVar and de novo variants in DDD cases versus controls

	Grantham score	Protein surface-exposure	Sequence conservation
ClinVar pathogenic variants	91.1	0.53	0.87
ClinVar benign variants	67.4	0.41	0.54
Difference in human expert annotations	+23.7	+0.12	+0.33
De novo variants in DDD patients	84.9	0.51	0.90
De novo variants in healthy controls	72.7	0.29	0.73
Difference in affected versus unaffected individuals	+12.2	+0.22	+0.17

Mean scores are shown for missense mutations with non-conflicting annotations in the ClinVar database, and for de novo variants present in DDD cases versus controls within 605 disease-associated genes. Protein surface-exposure reflects the fraction of amino acids predicted as exposed residues by the solvent accessibility neural network, and sequence conservation shows the fraction of amino acids with sequence identity in the 100-vertebrate alignment. Numbers in bold highlight differences in the heuristics favored by human experts compared to empirical data.

FIG. 63

Per-Gene Enrichment Analysis



$$\sum (observed\ trinucleotide\ mutation\ rates) = 10^{-5}$$

Transmission (Chromosome) Count

Baseline Number of Mutations = $2 \times 1000 \times 10^{-5} = 0.2$

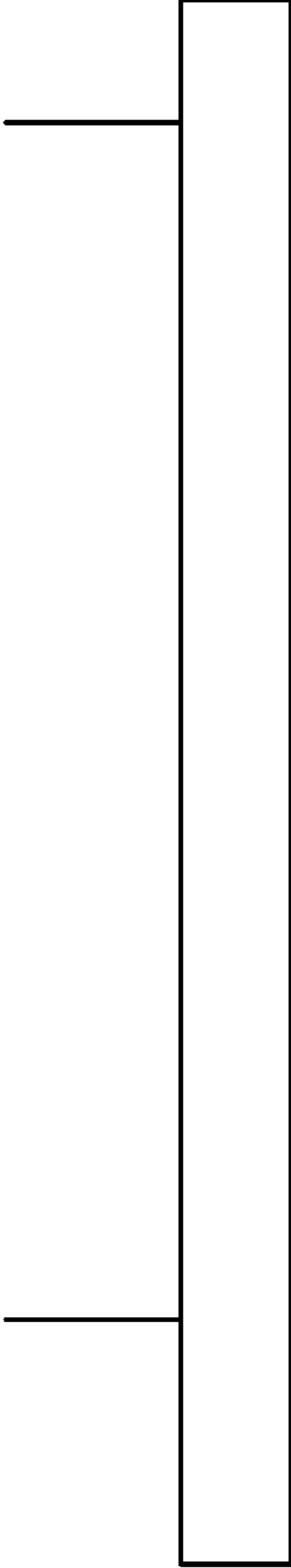
Cohort Size

De Novo Variant Count = 3

FIG. 64

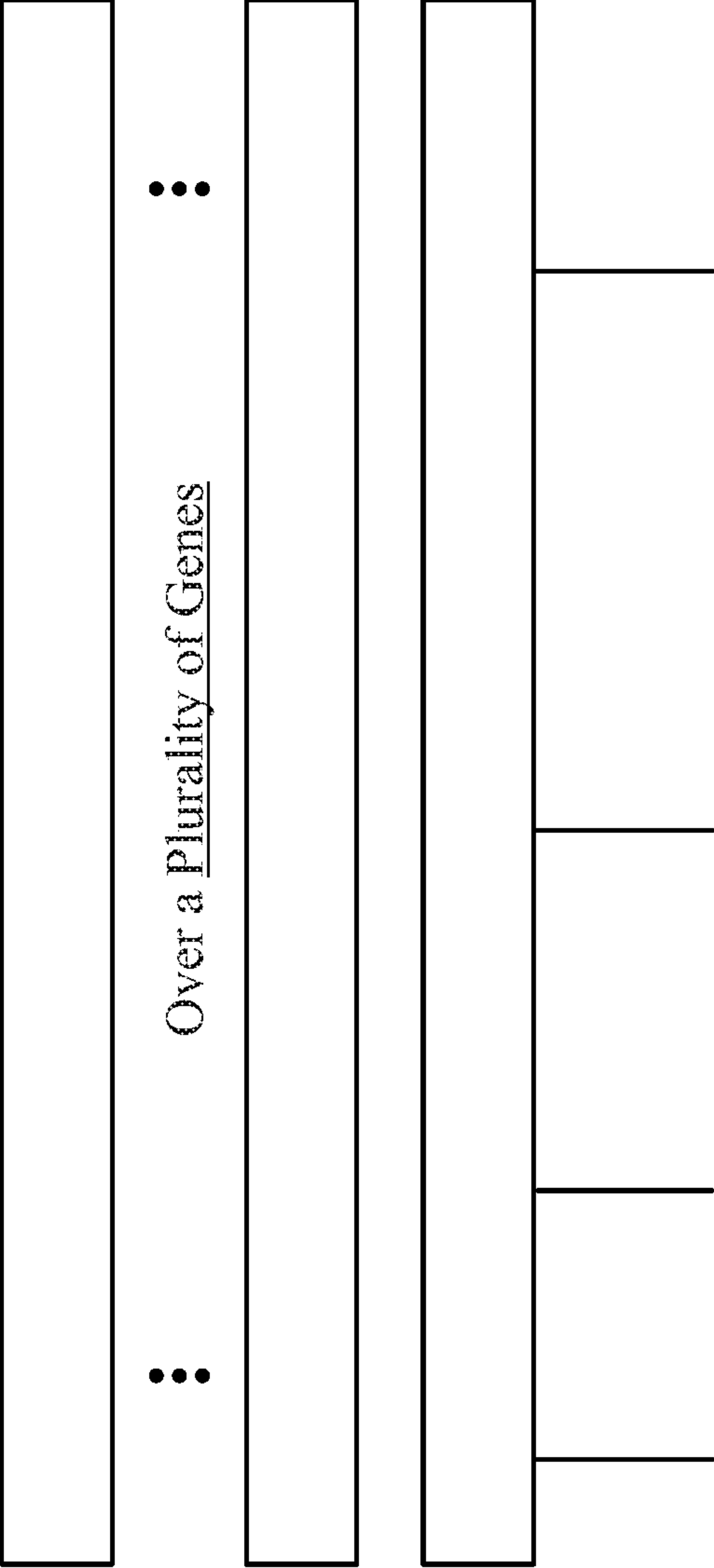
Genome-Wide Enrichment Analysis

Cohort of 2000 Healthy Individuals
Classified as Pathogenic



⋮ Over a Plurality of Genes ⋮

Cohort of 1000 Affected Individuals
Classified as Causing Aberrant Splicing



$$\text{Mutation Rate in Healthy Cohort} = 2/2000 = 0.001$$

$$\text{Mutation Rate in Affected Cohort} = 4/1000 = 0.004$$

$$\text{Per-Individual Mutation Ratio} = 0.004/0.001 = 4$$

FIG. 65

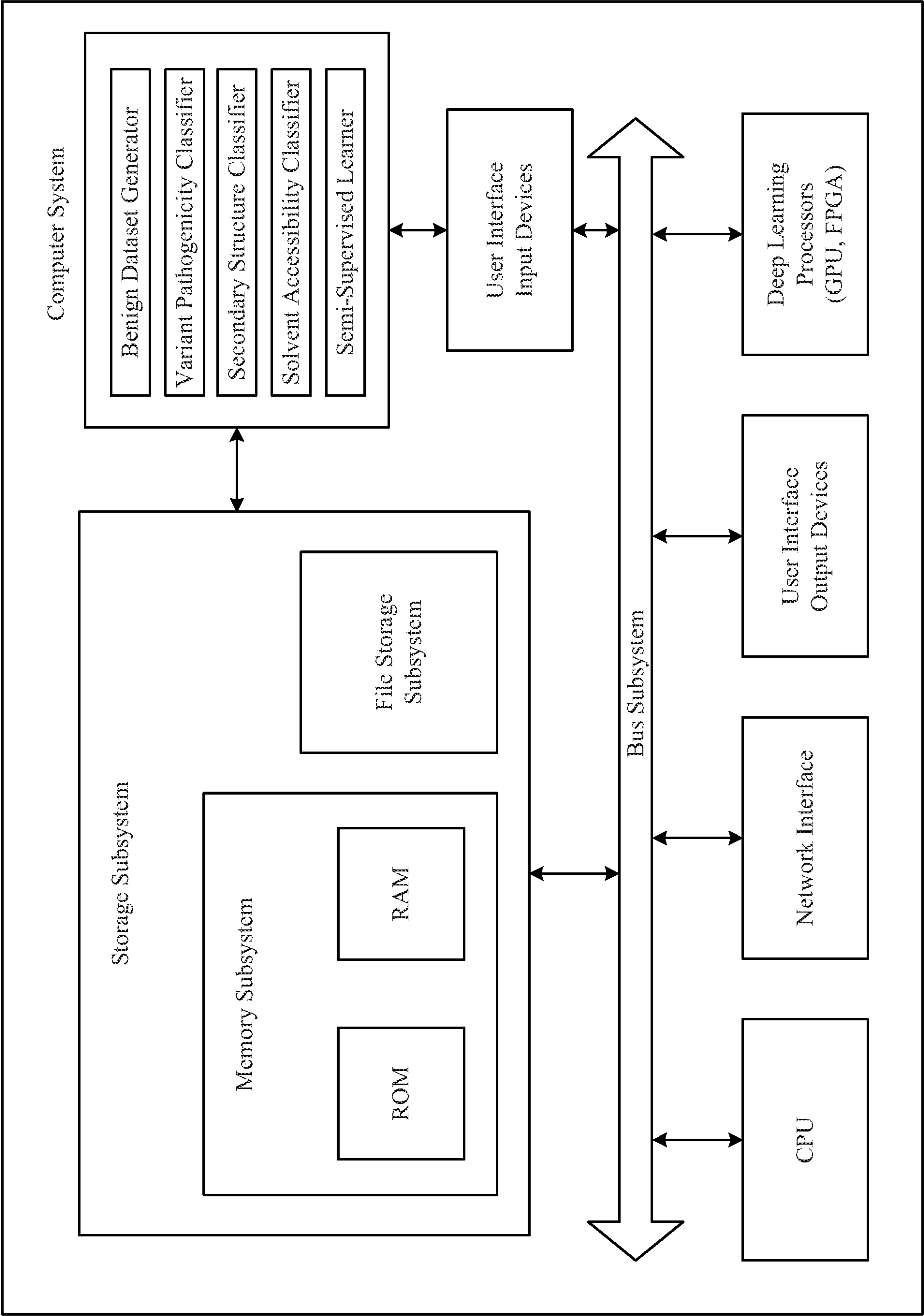


FIG. 66