

In the name of the most high

Introduction to Bioinformatics

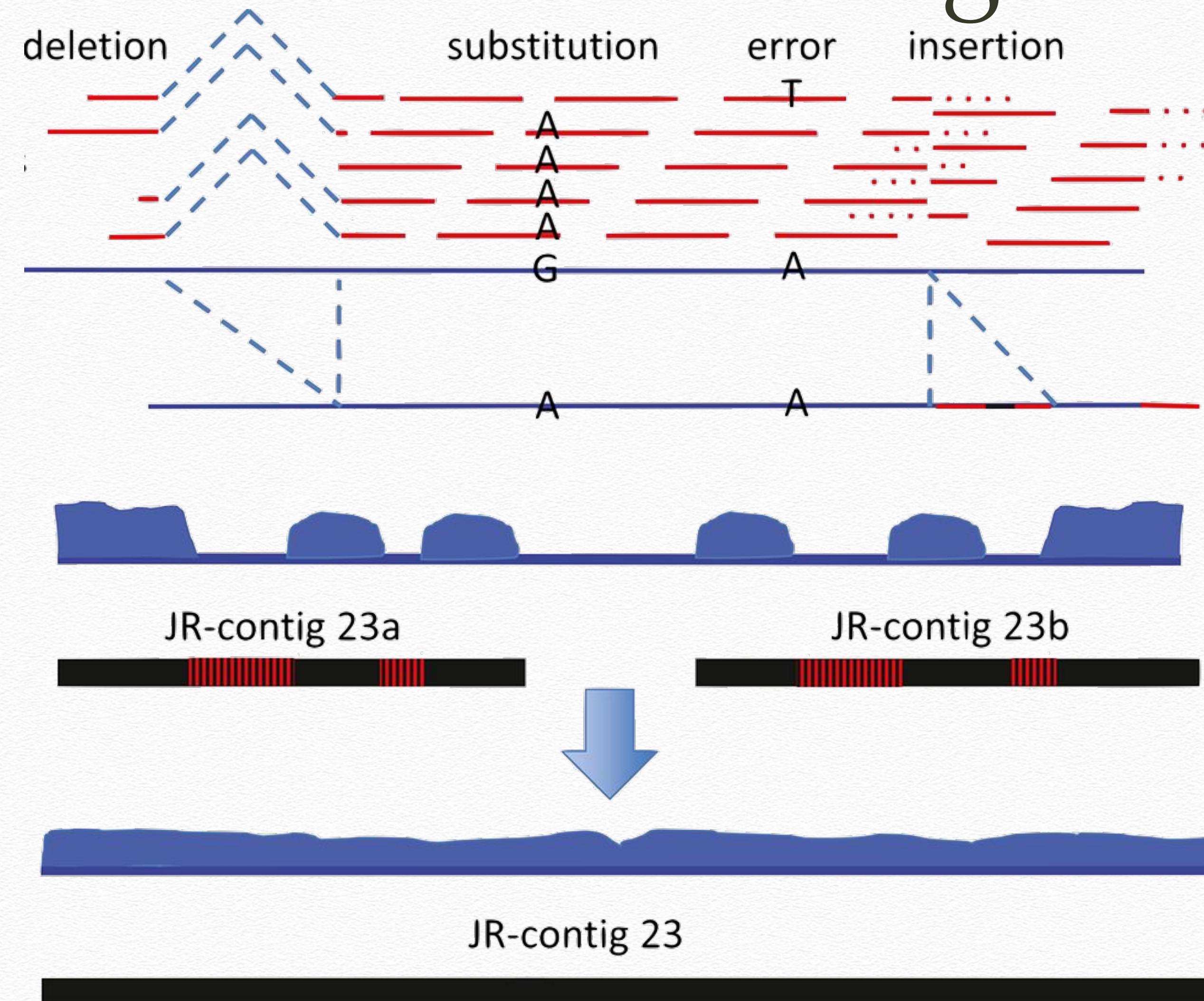
Genotyping

Ali Sharifi-Zarchi

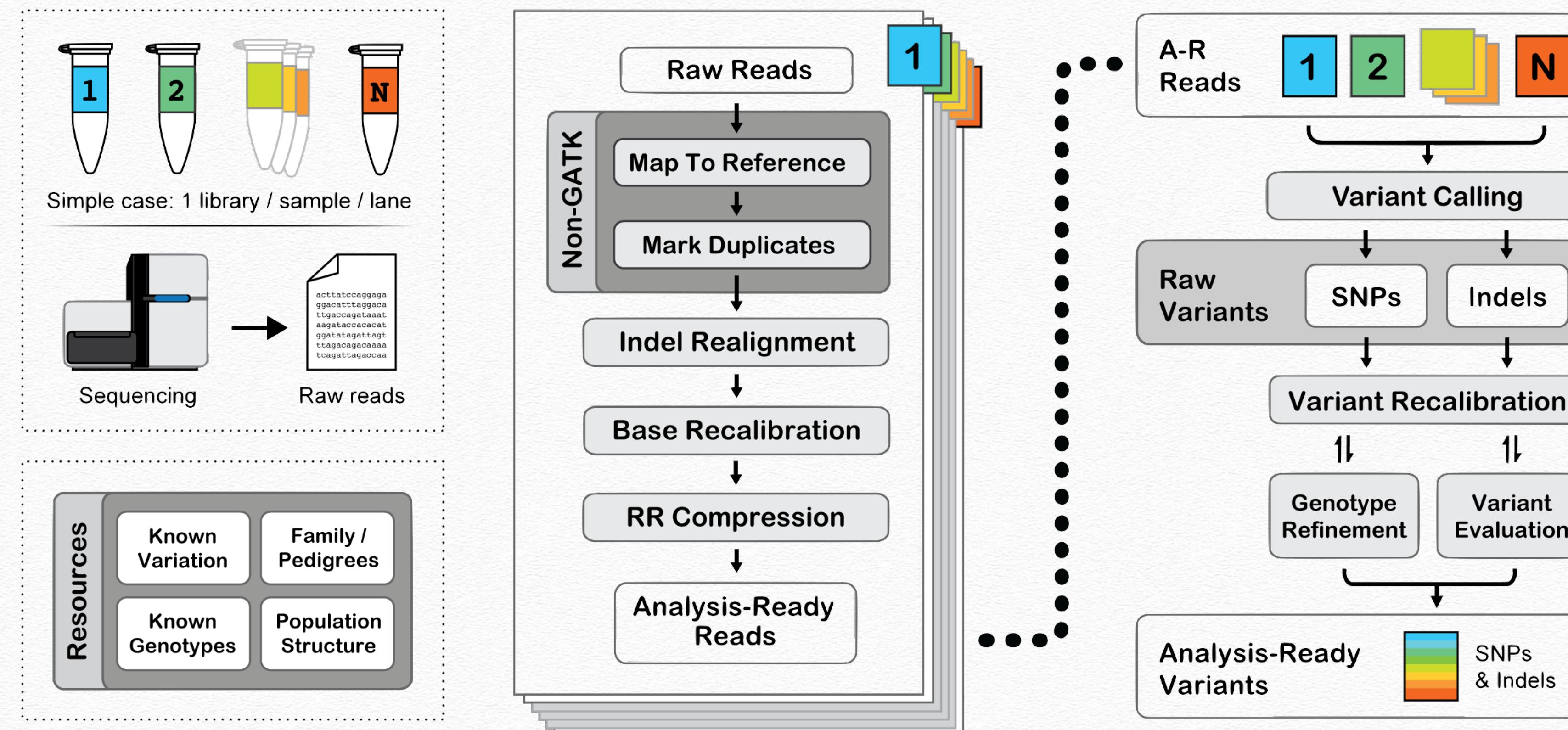
Department of Computer Engineering, Sharif University of Technology

These slides are available under the Creative Commons Attribution License.

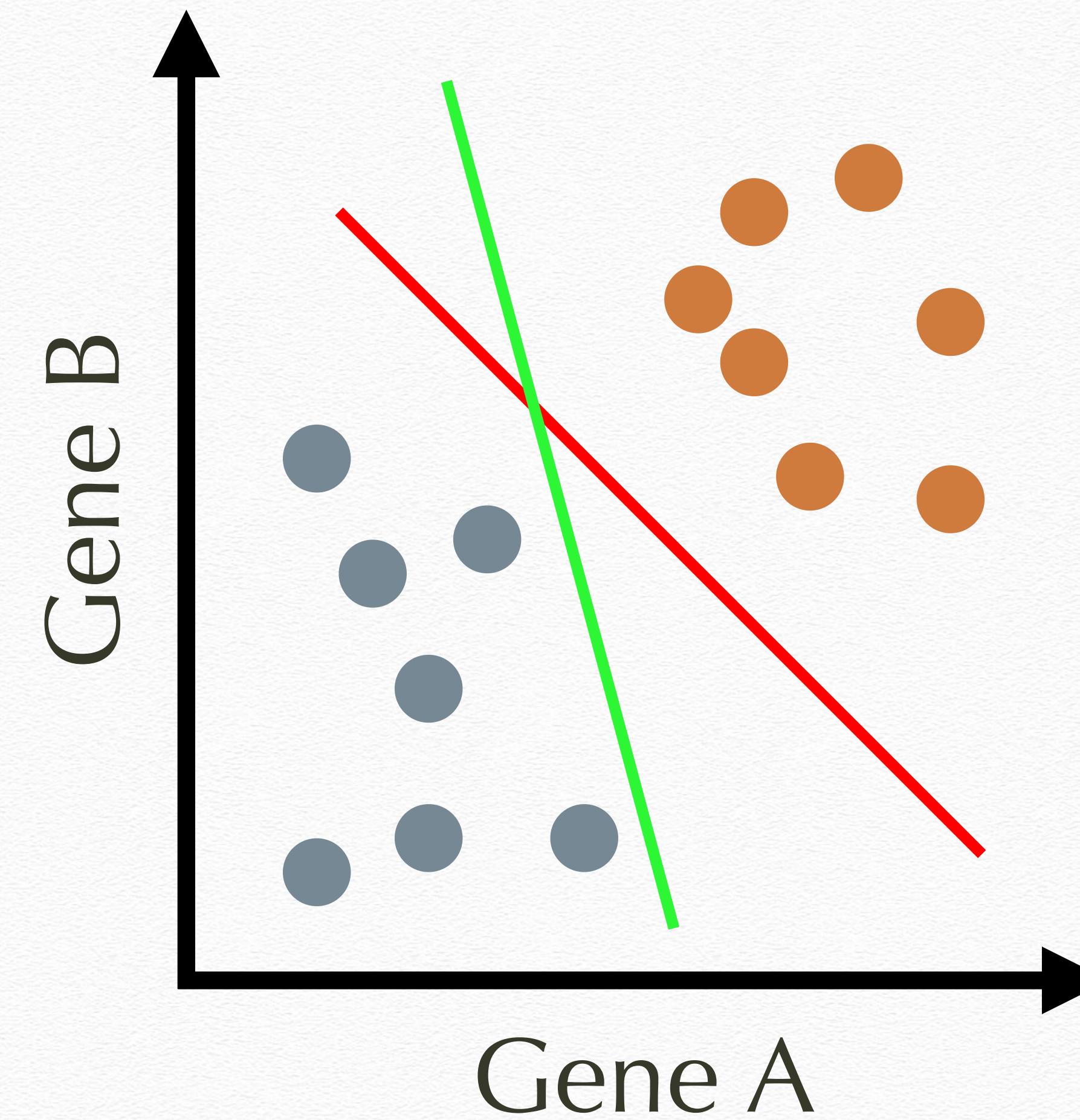
Variant Calling



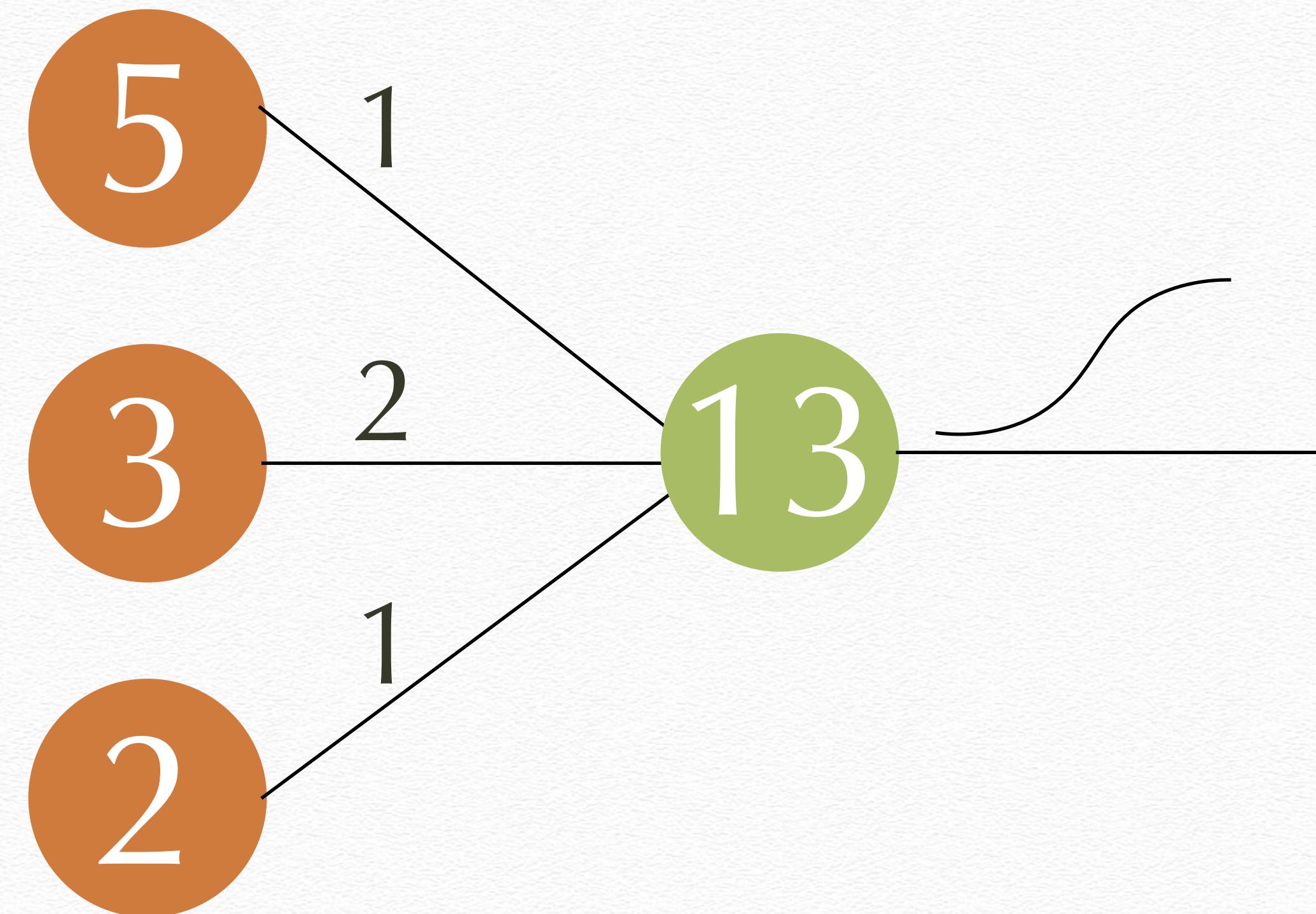
GATK



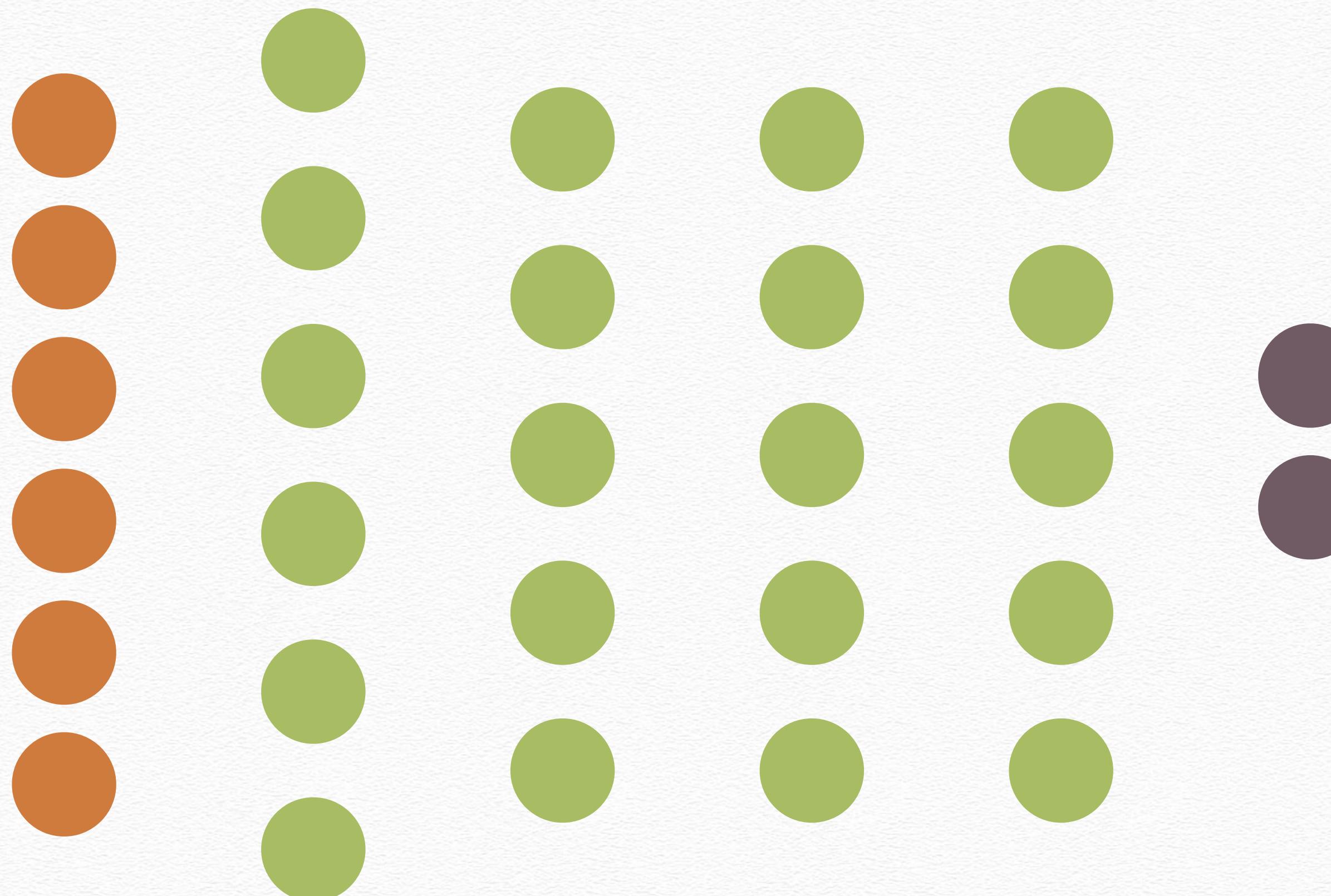
Classification



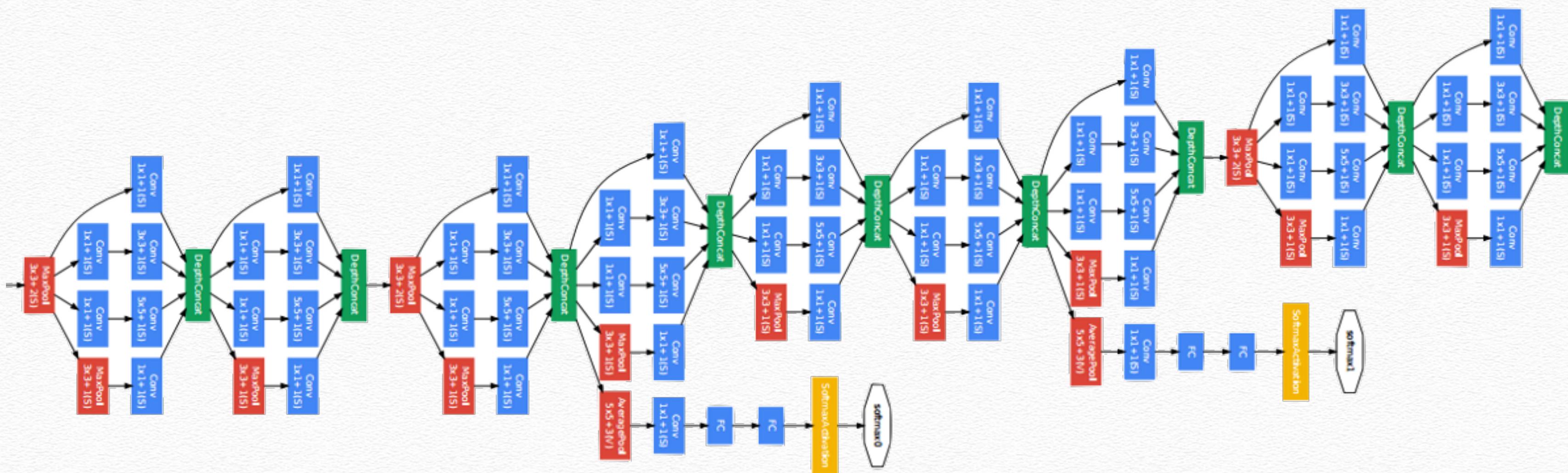
Neural Networks



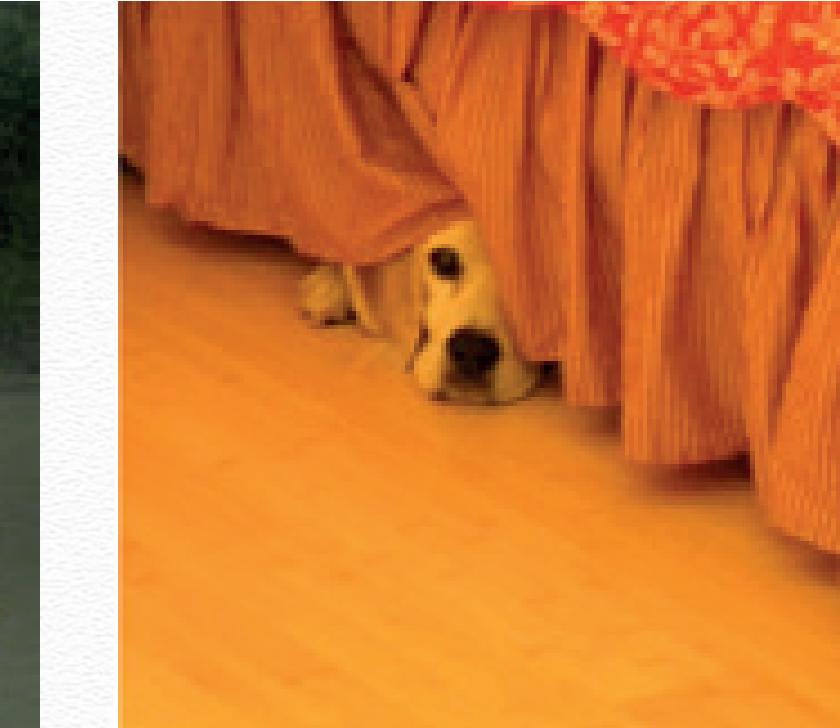
Deep Neural Networks



Inception

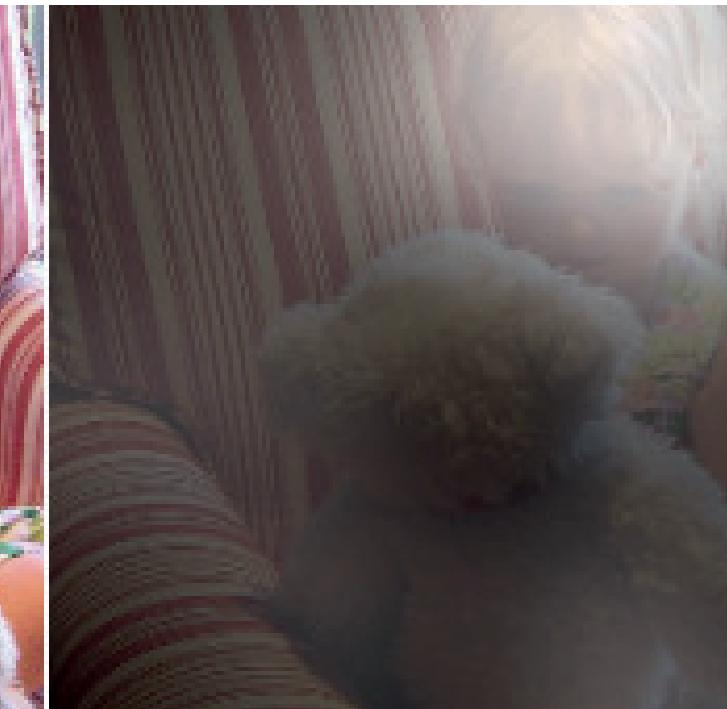


From Image to Text



A woman is throwing a **frisbee** in a park.

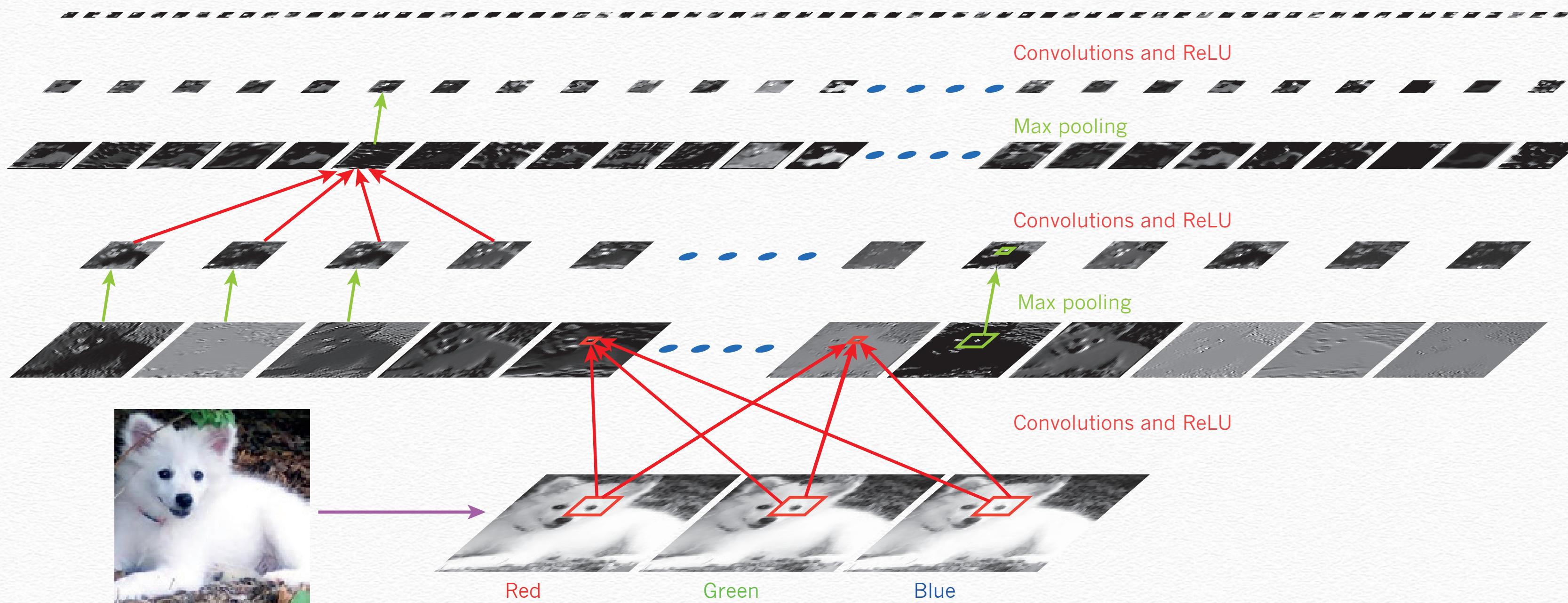
A **dog** is standing on a hardwood floor.



A little **girl** sitting on a bed with a teddy bear.

A group of **people** sitting on a boat in the water.

Features



Even Games!

LETTER

doi:10.1038/nature14236

Human-level control through deep reinforcement learning

Volodymyr Mnih^{1*}, Koray Kavukcuoglu^{1*}, David Silver^{1*}, Andrei A. Rusu¹, Joel Veness¹, Marc G. Bellemare¹, Alex Graves¹, Martin Riedmiller¹, Andreas K. Fidjeland¹, Georg Ostrovski¹, Stig Petersen¹, Charles Beattie¹, Amir Sadik¹, Ioannis Antonoglou¹, Helen King¹, Dharshan Kumaran¹, Daan Wierstra¹, Shane Legg¹ & Demis Hassabis¹

The theory of reinforcement learning provides a normative account¹, deeply rooted in psychological² and neuroscientific³ perspectives on animal behaviour, of how agents may optimize their control of an environment. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted with a difficult task: they must derive efficient representations of the environment from high-dimensional sensory inputs, and use these to generalize past experience to new situations. Remarkably, humans and other animals seem to solve this problem through a harmonious combination of reinforcement learning and hierarchical sensory processing systems^{4,5}, the former evidenced by a wealth of neural data revealing notable parallels between the phasic signals emitted by dopaminergic neurons and temporal difference reinforcement learning algorithms³. While reinforcement learning agents have achieved some successes in a variety of domains^{6–8}, their applicability has previously been limited to domains in which useful features can be handcrafted,

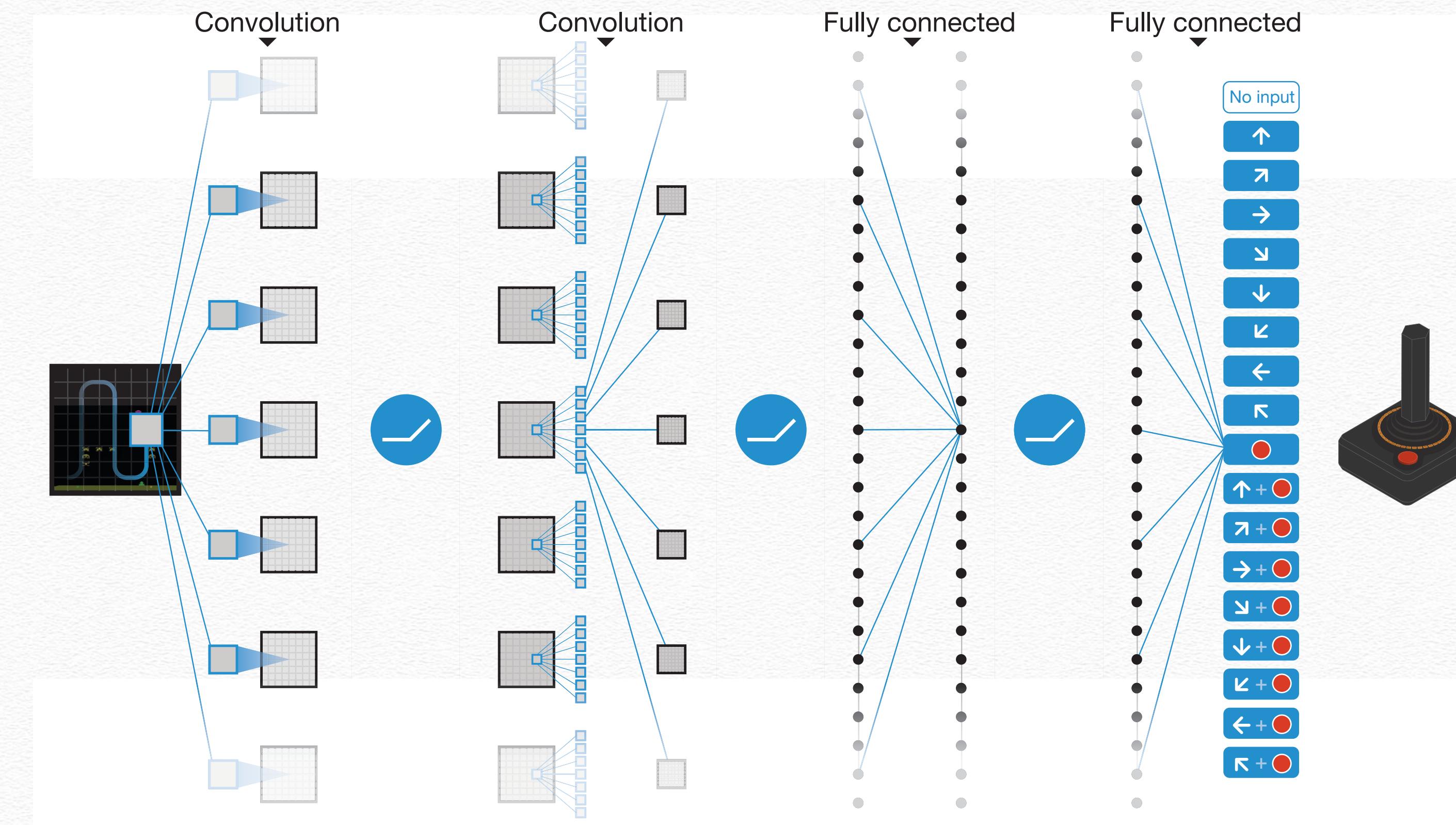
agent is to select actions in a fashion that maximizes cumulative future reward. More formally, we use a deep convolutional neural network to approximate the optimal action-value function

$$Q^*(s, a) = \max_{\pi} \mathbb{E}[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, a_t = a, \pi],$$

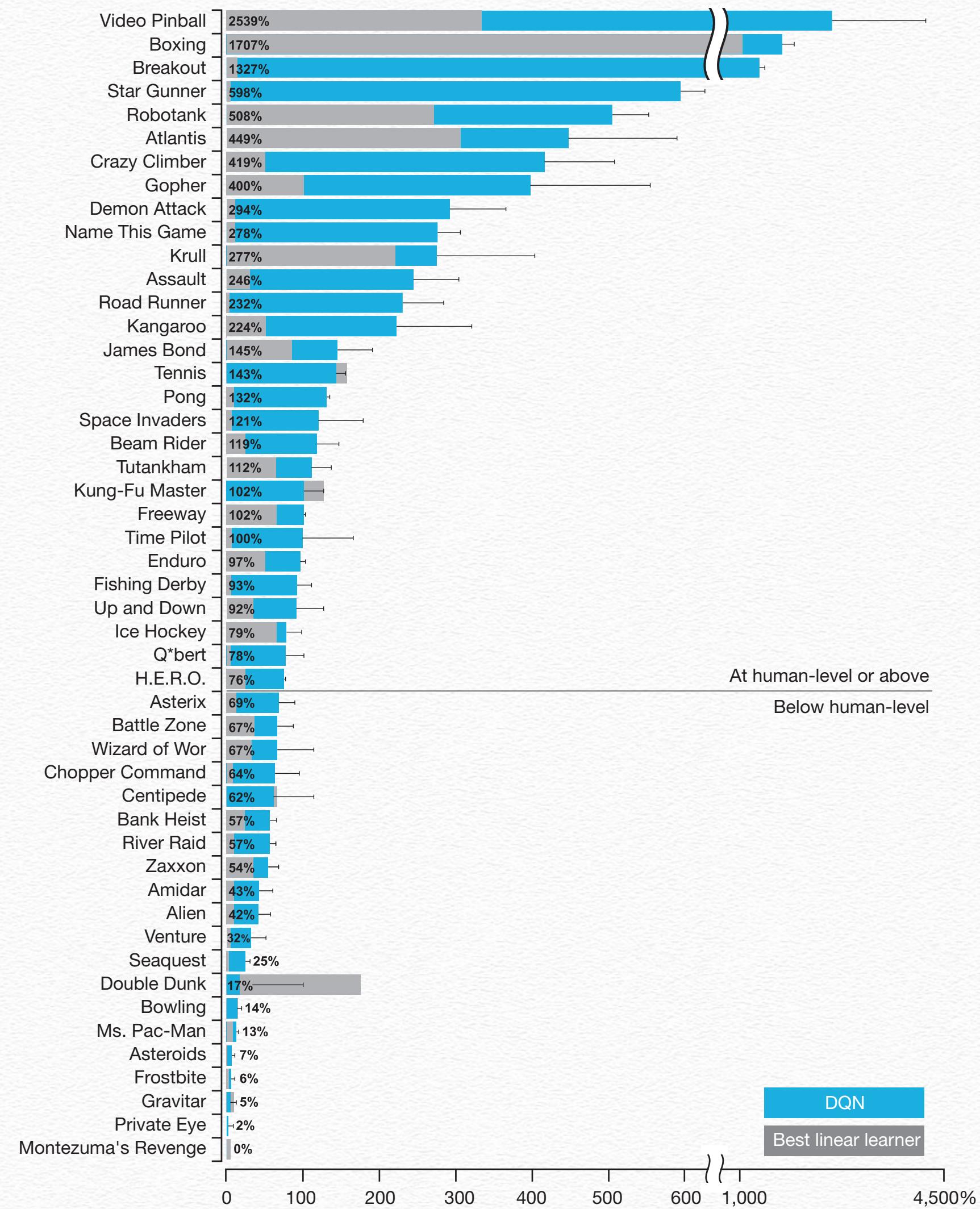
which is the maximum sum of rewards r_t discounted by γ at each time-step t , achievable by a behaviour policy $\pi = P(a|s)$, after making an observation (s) and taking an action (a) (see Methods)¹⁹.

Reinforcement learning is known to be unstable or even to diverge when a nonlinear function approximator such as a neural network is used to represent the action-value (also known as Q) function²⁰. This instability has several causes: the correlations present in the sequence of observations, the fact that small updates to Q may significantly change the policy and therefore change the data distribution, and the correlations between the action-values (Q) and the target values $r + \gamma \max_{a'} Q(s', a')$. We address these instabilities with a novel variant of Q-learning, which

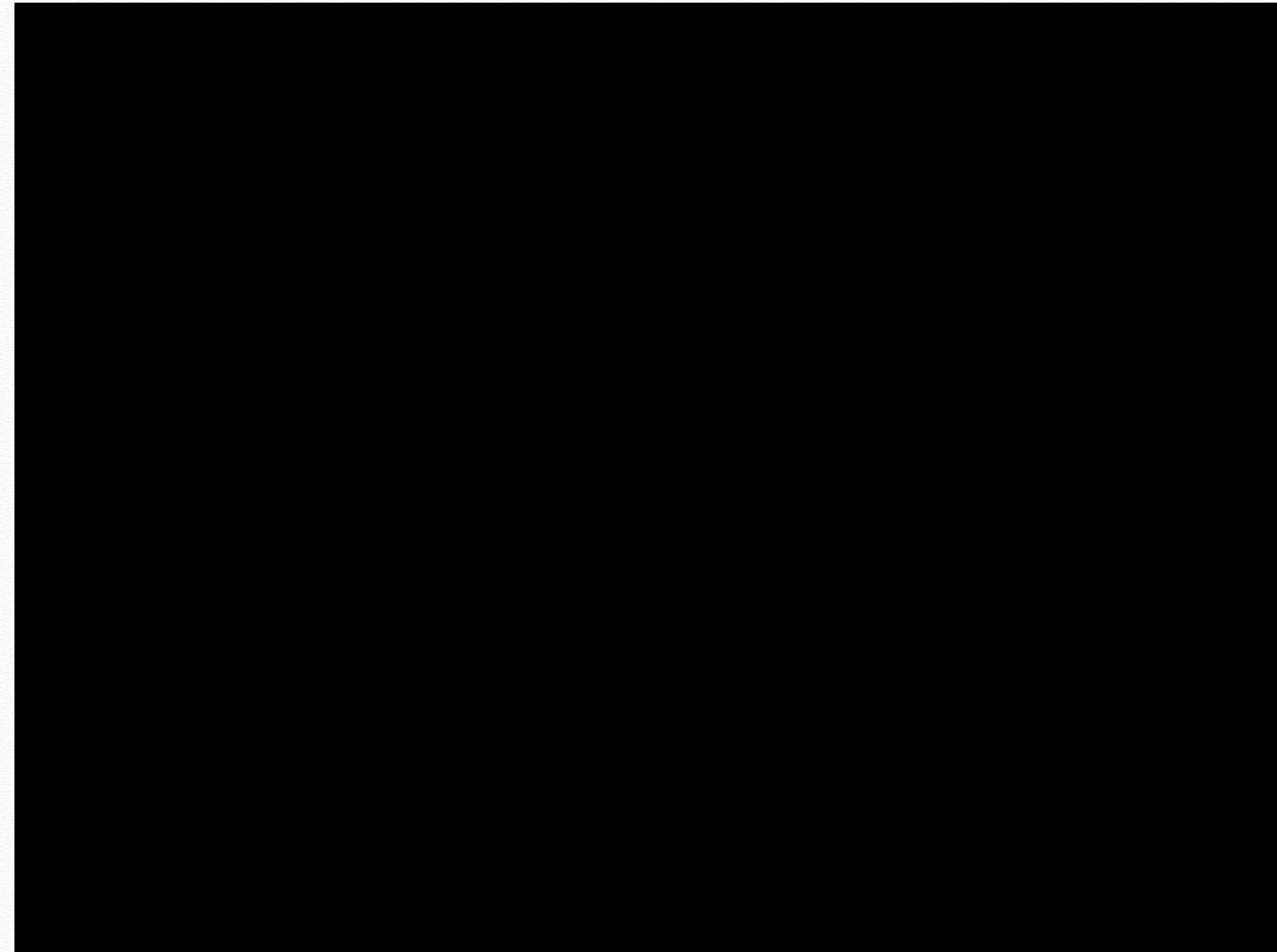
Even Games!



Even Games!



Even Games!



GO

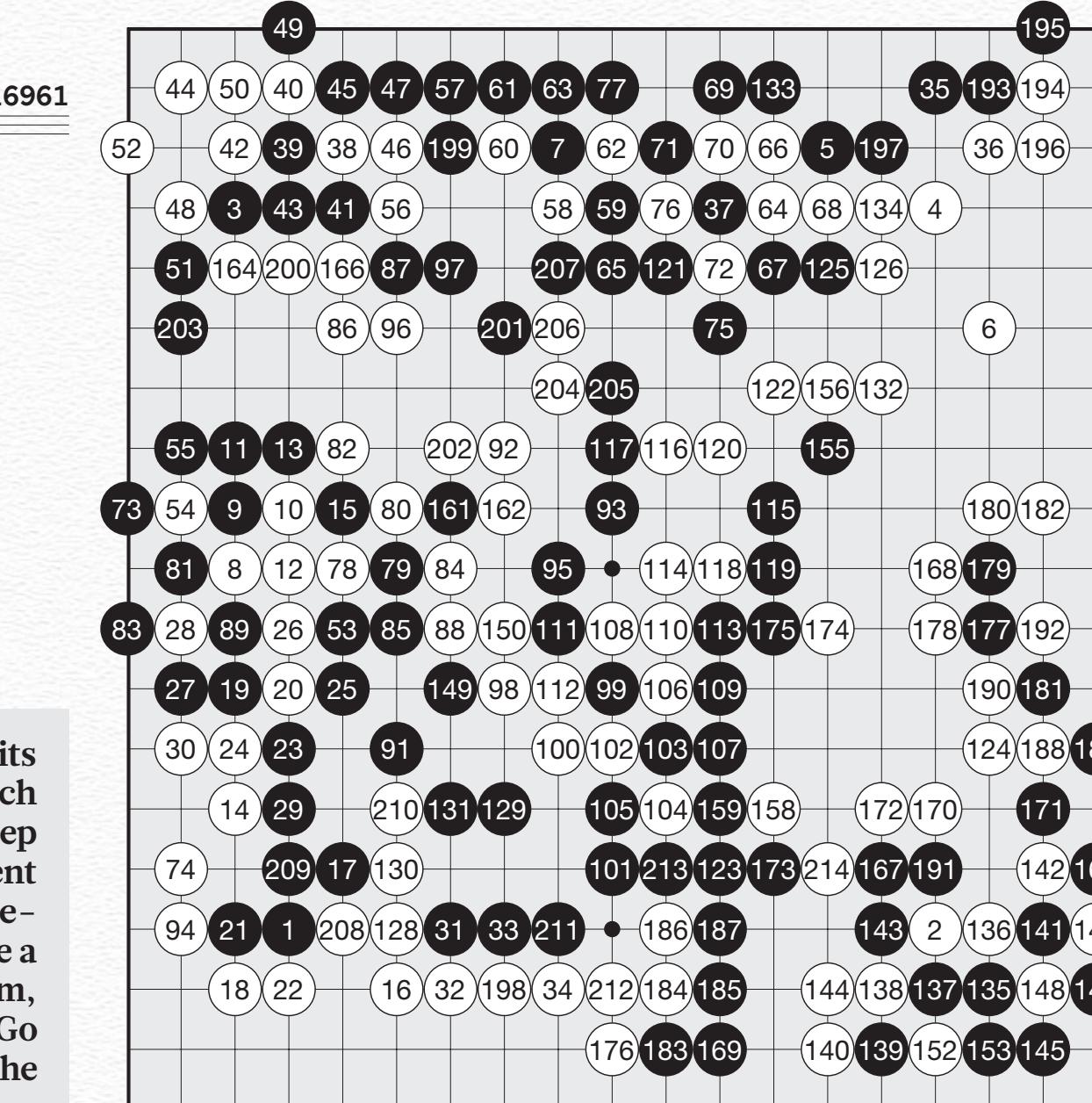
ARTICLE

doi:10.1038/nature16961

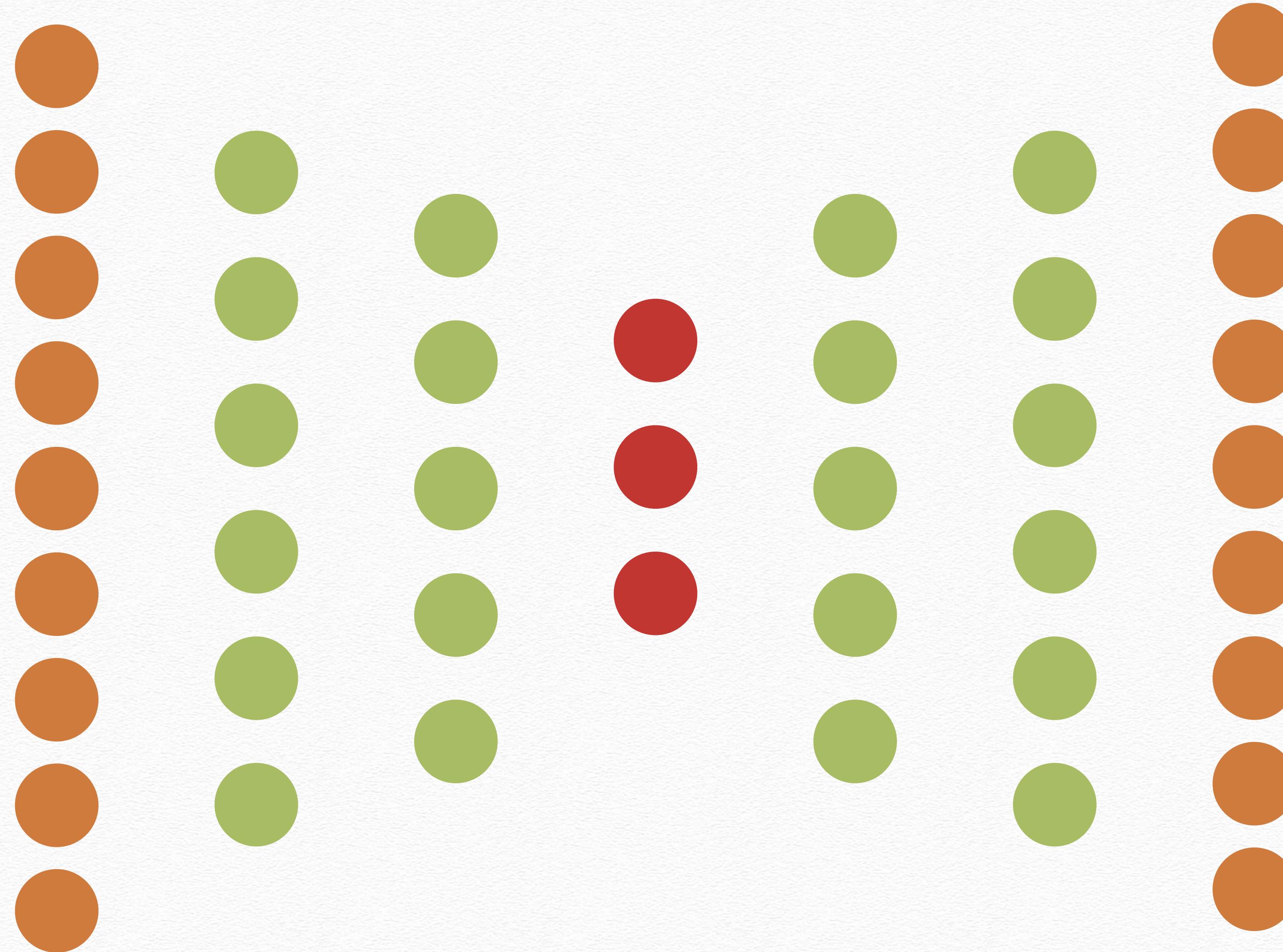
Mastering the game of Go with deep neural networks and tree search

David Silver^{1*}, Aja Huang^{1*}, Chris J. Maddison¹, Arthur Guez¹, Laurent Sifre¹, George van den Driessche¹, Julian Schrittwieser¹, Ioannis Antonoglou¹, Veda Panneershelvam¹, Marc Lanctot¹, Sander Dieleman¹, Dominik Grewe¹, John Nham², Nal Kalchbrenner¹, Ilya Sutskever², Timothy Lillicrap¹, Madeleine Leach¹, Koray Kavukcuoglu¹, Thore Graepel¹ & Demis Hassabis¹

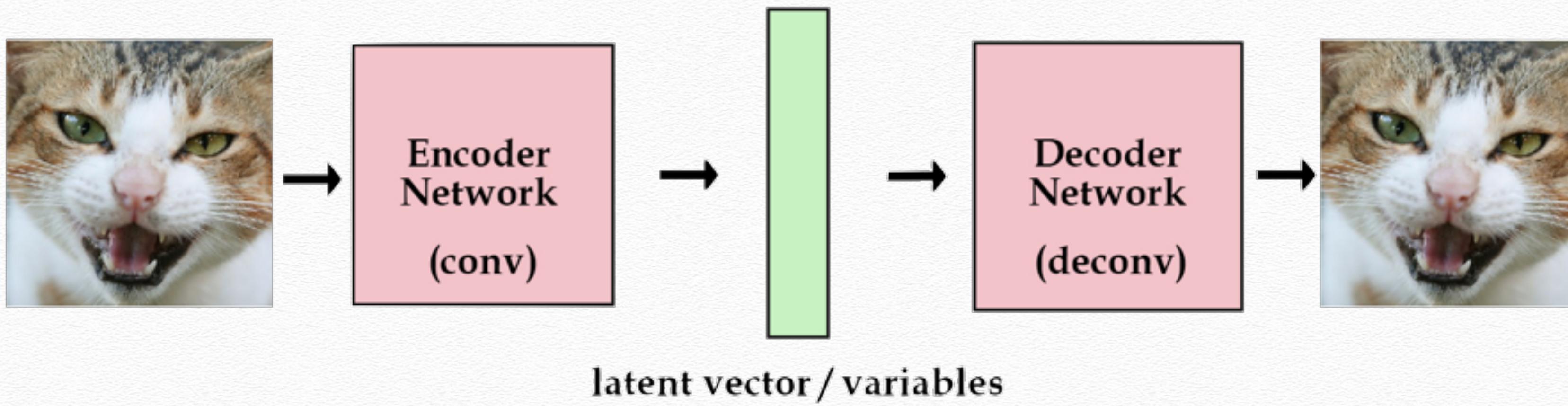
The game of Go has long been viewed as the most challenging of classic games for artificial intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves. Here we introduce a new approach to computer Go that uses ‘value networks’ to evaluate board positions and ‘policy networks’ to select moves. These deep neural networks are trained by a novel combination of supervised learning from human expert games, and reinforcement learning from games of self-play. Without any lookahead search, the neural networks play Go at the level of state-of-the-art Monte Carlo tree search programs that simulate thousands of random games of self-play. We also introduce a new search algorithm that combines Monte Carlo simulation with value and policy networks. Using this search algorithm, our program AlphaGo achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion by 5 games to 0. This is the first time that a computer program has defeated a human professional player in the full-sized game of Go, a feat previously thought to be at least a decade away.



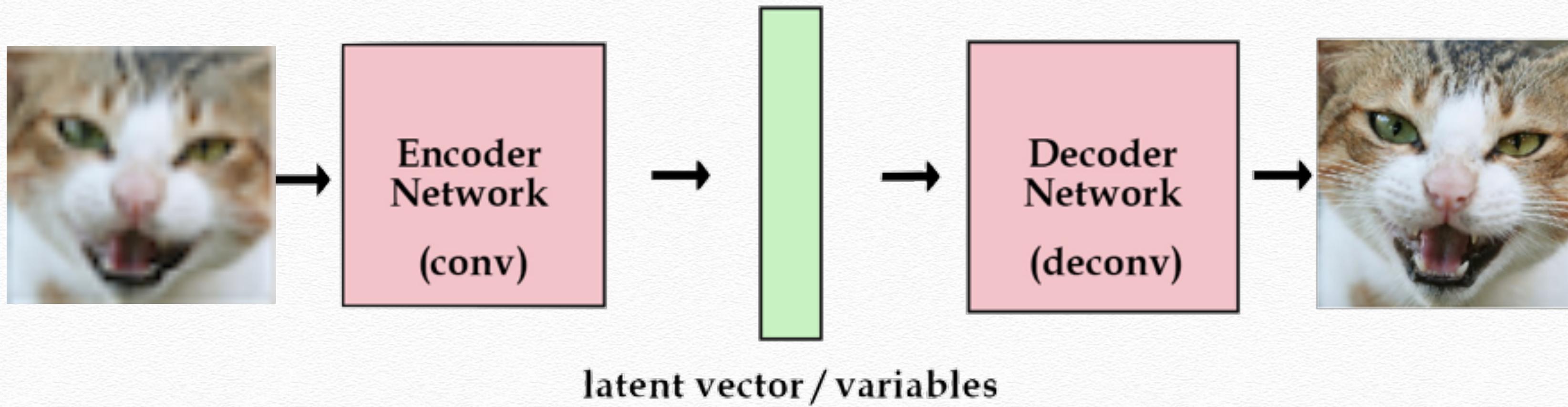
Autoencoders



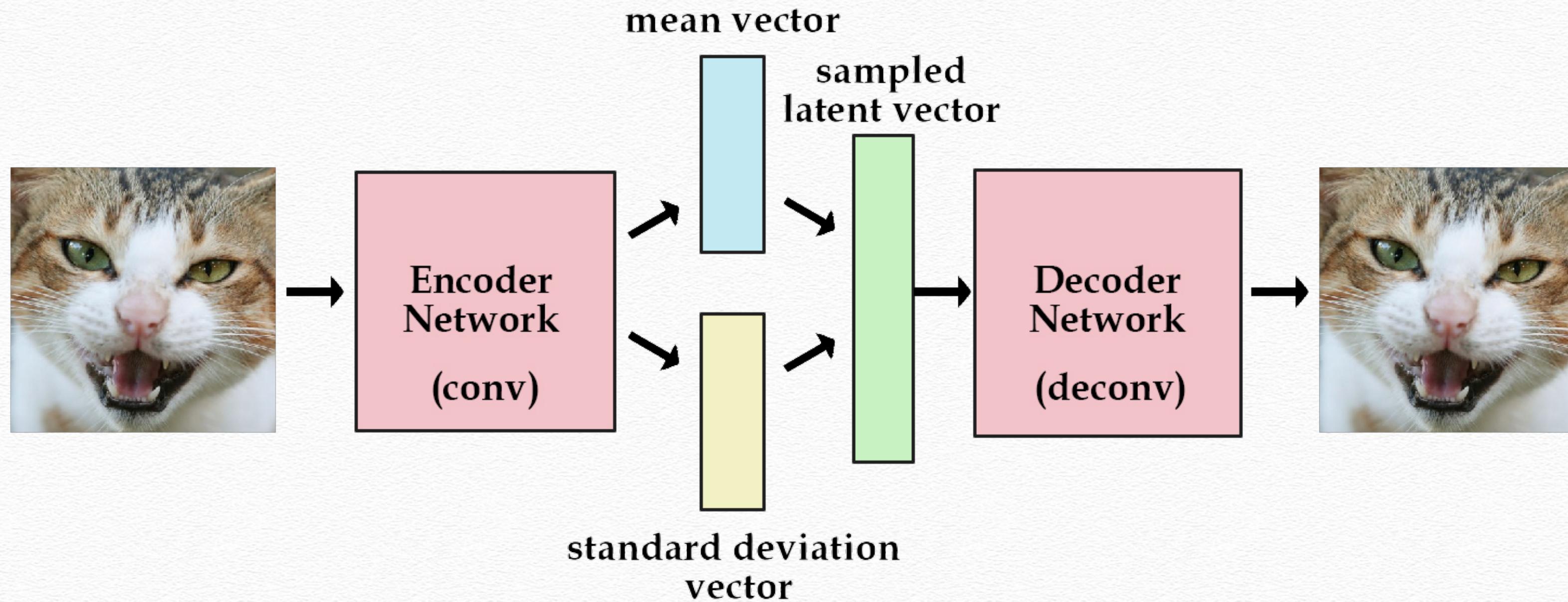
Autoencoders



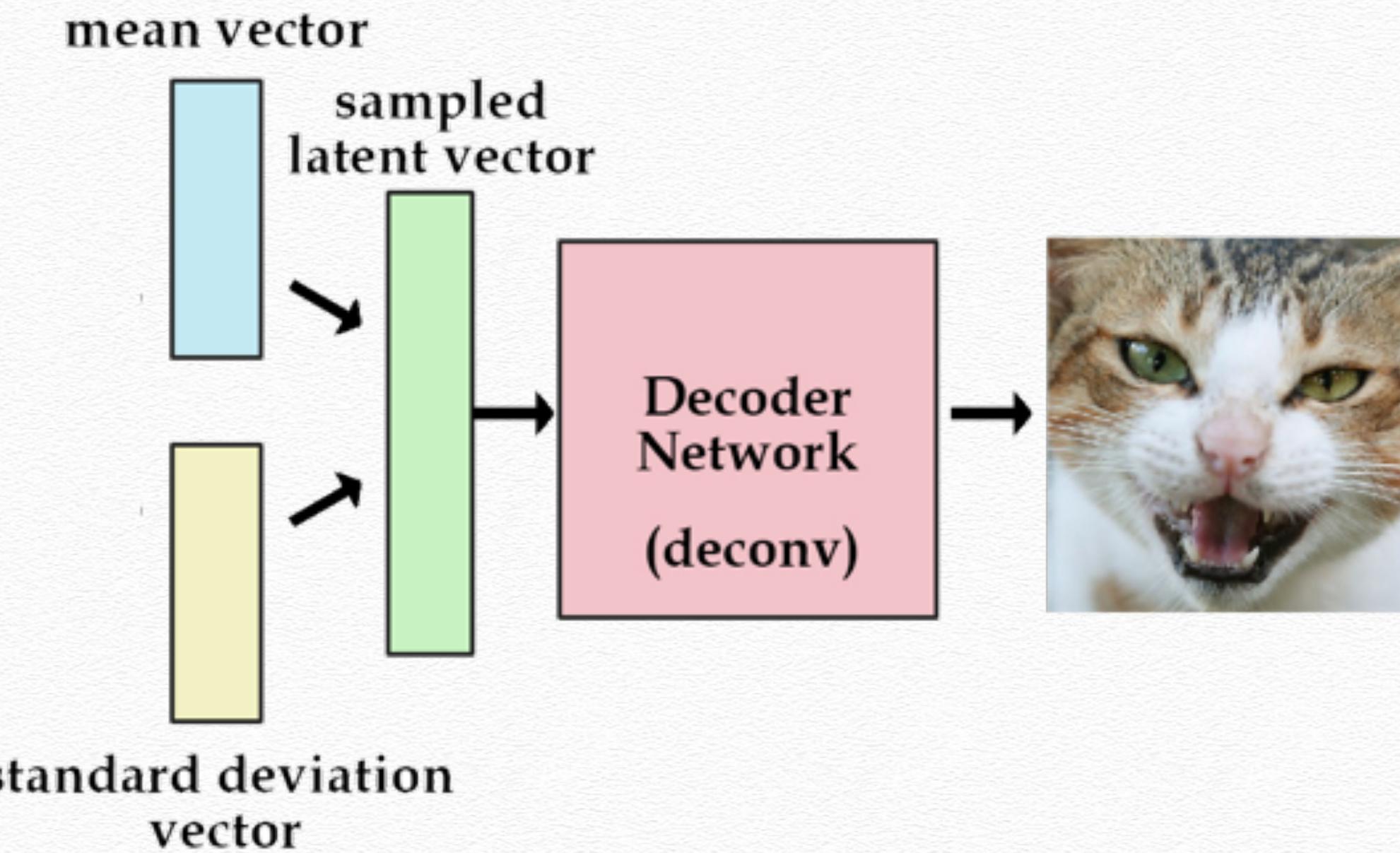
Denoising Autoencoders



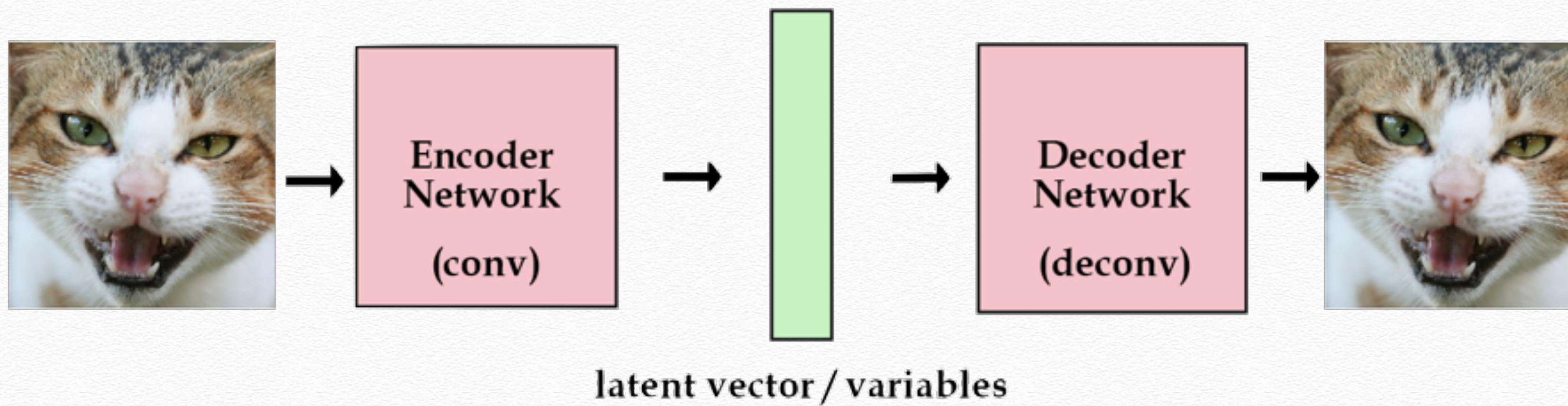
Variational Autoencoders



Variational Autoencoders



Contractive Autoencoders



$$\Omega(h) = \lambda \left\| \frac{\partial(f(x))}{\partial(x)} \right\|_F^2$$

THE NEURAL NETWORK ZOO

A mostly complete chart of

Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

Backfed Input Cell

Input Cell

Noisy Input Cell

Hidden Cell

Probabilistic Hidden Cell

Spiking Hidden Cell

Output Cell

Match Input Output Cell

Recurrent Cell

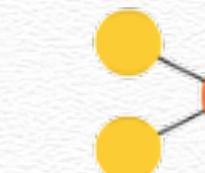
Memory Cell

Different Memory Cell

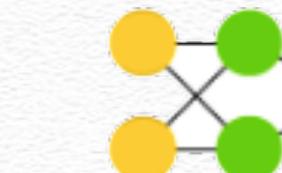
Kernel

Convolution or Pool

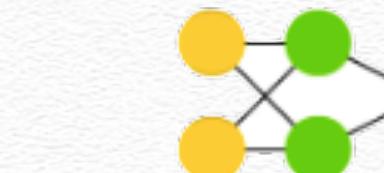
Perceptron (P)



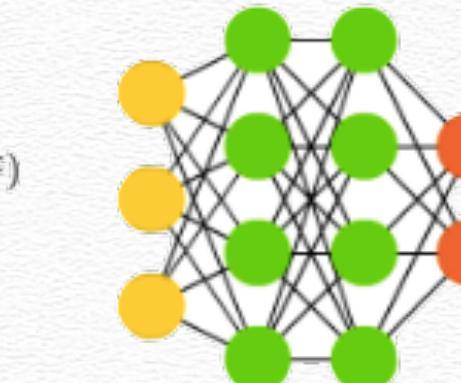
Feed Forward (FF)



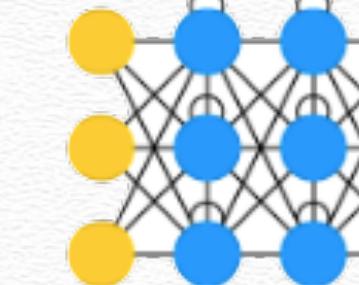
Radial Basis Network (RBF)



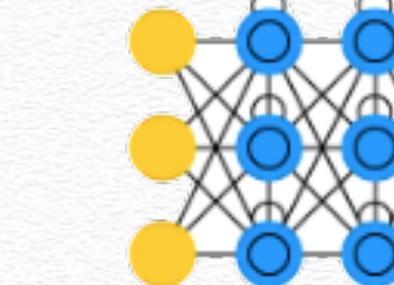
Deep Feed Forward (DFF)



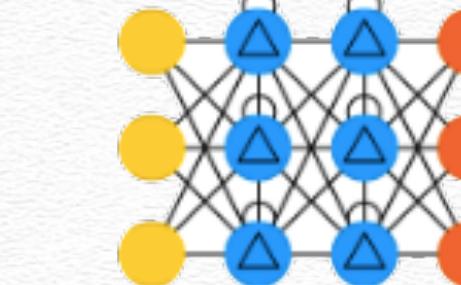
Recurrent Neural Network (RNN)



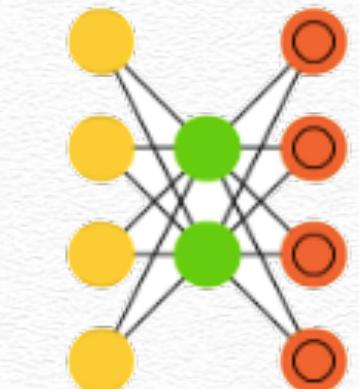
Long / Short Term Memory (LSTM)



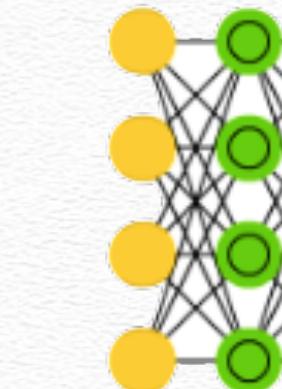
Gated Recurrent Unit (GRU)



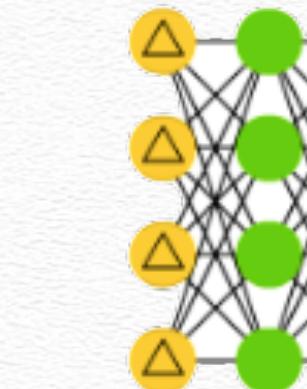
Auto Encoder (AE)



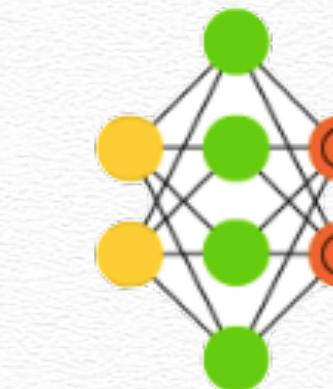
Variational AE (VAE)



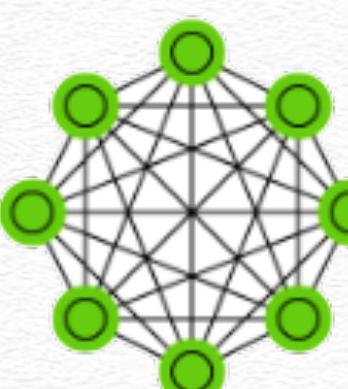
Denoising AE (DAE)



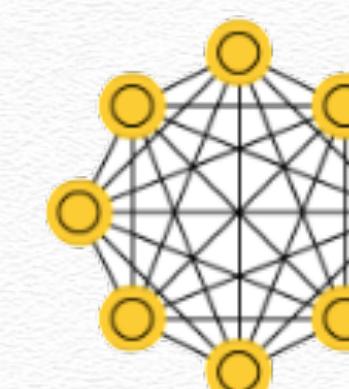
Sparse AE (SAE)



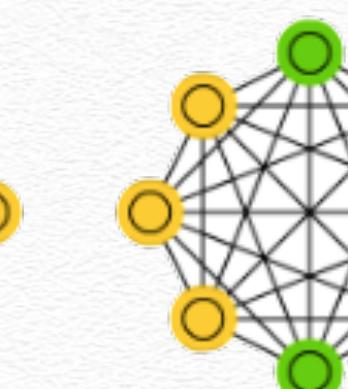
Markov Chain (MC)



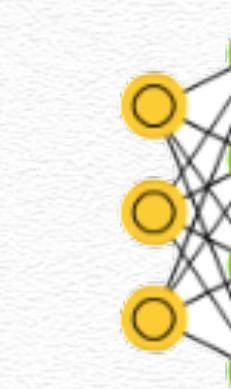
Hopfield Network (HN)



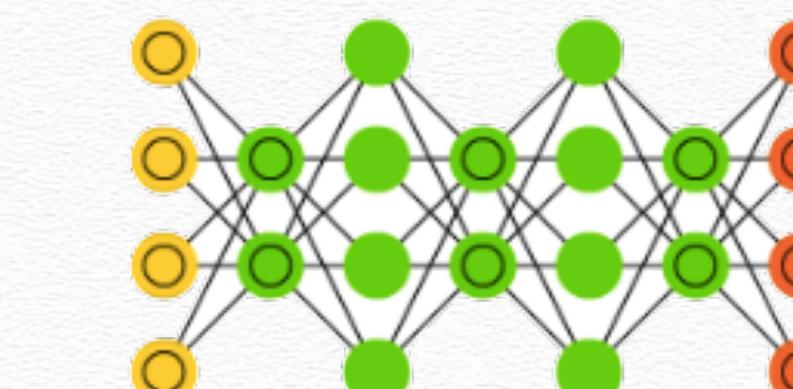
Boltzmann Machine (BM)



Restricted BM (RBM)



Deep Belief Network (DBN)



Deep Convolutional Network (DCN)

Deconvolutional Network (DN)

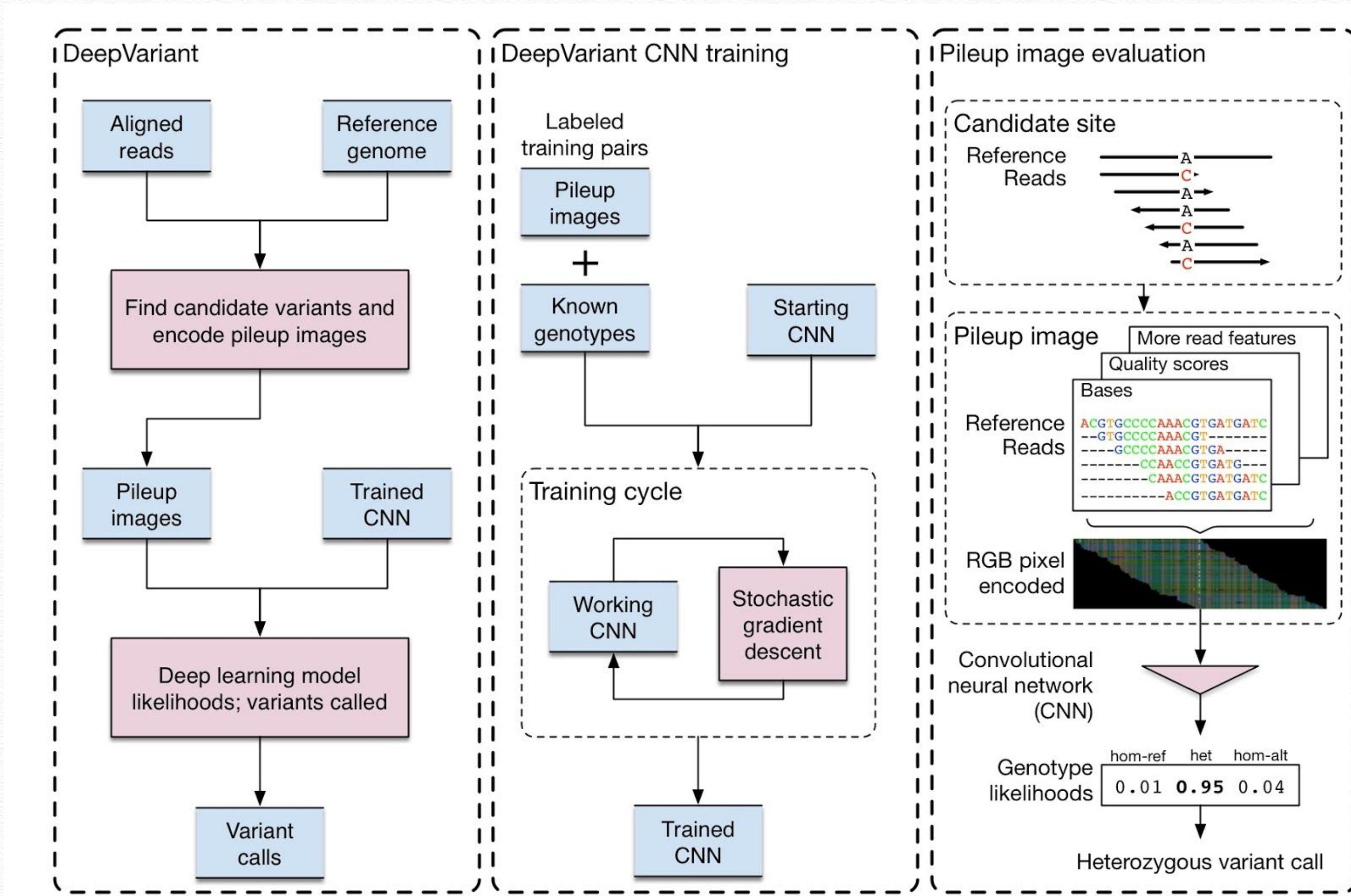
Deep Convolutional Inverse Graphics Network (DCIGN)

Variational Autoencoders

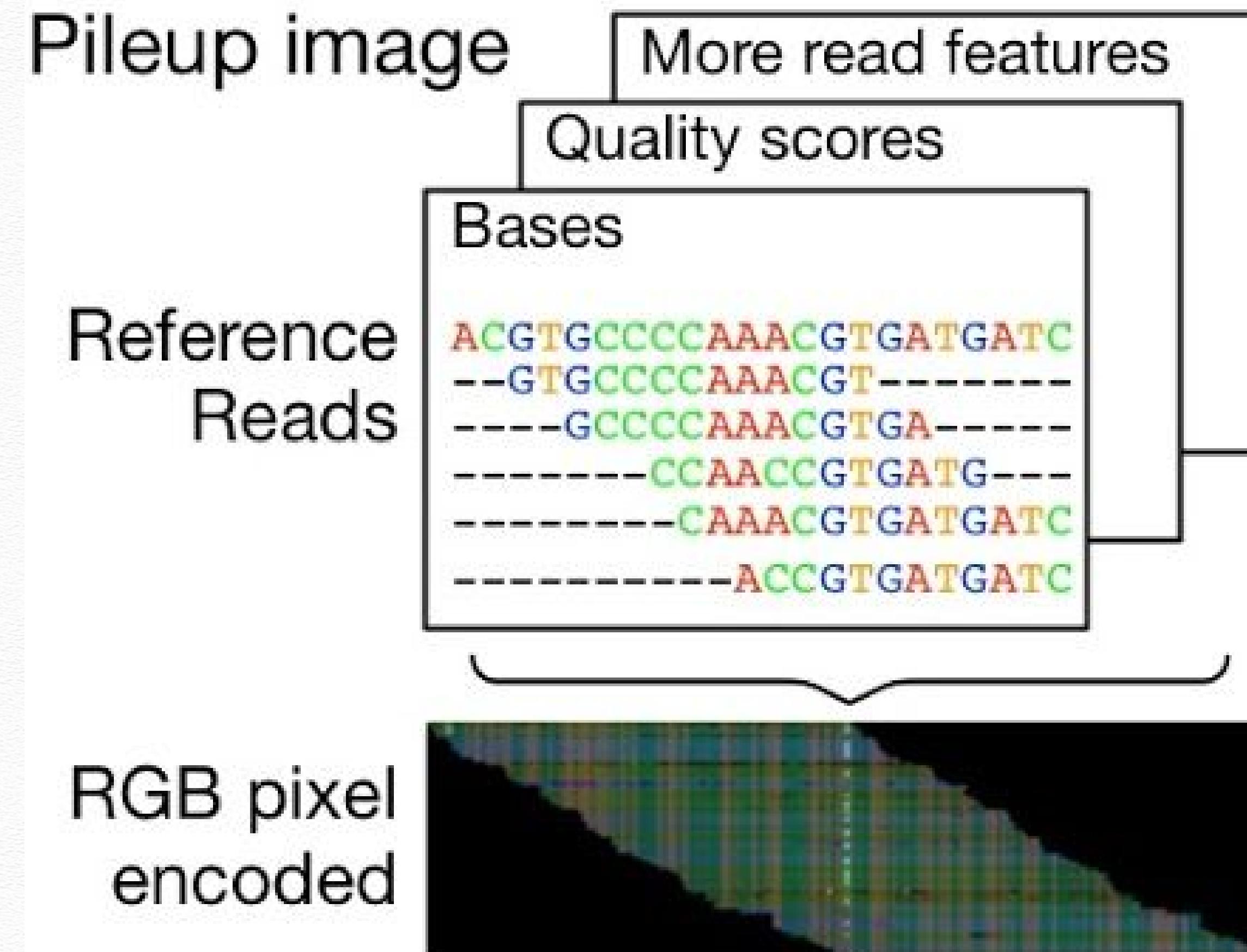


<https://jaan.io/what-is-variational-autoencoder-vae-tutorial/>

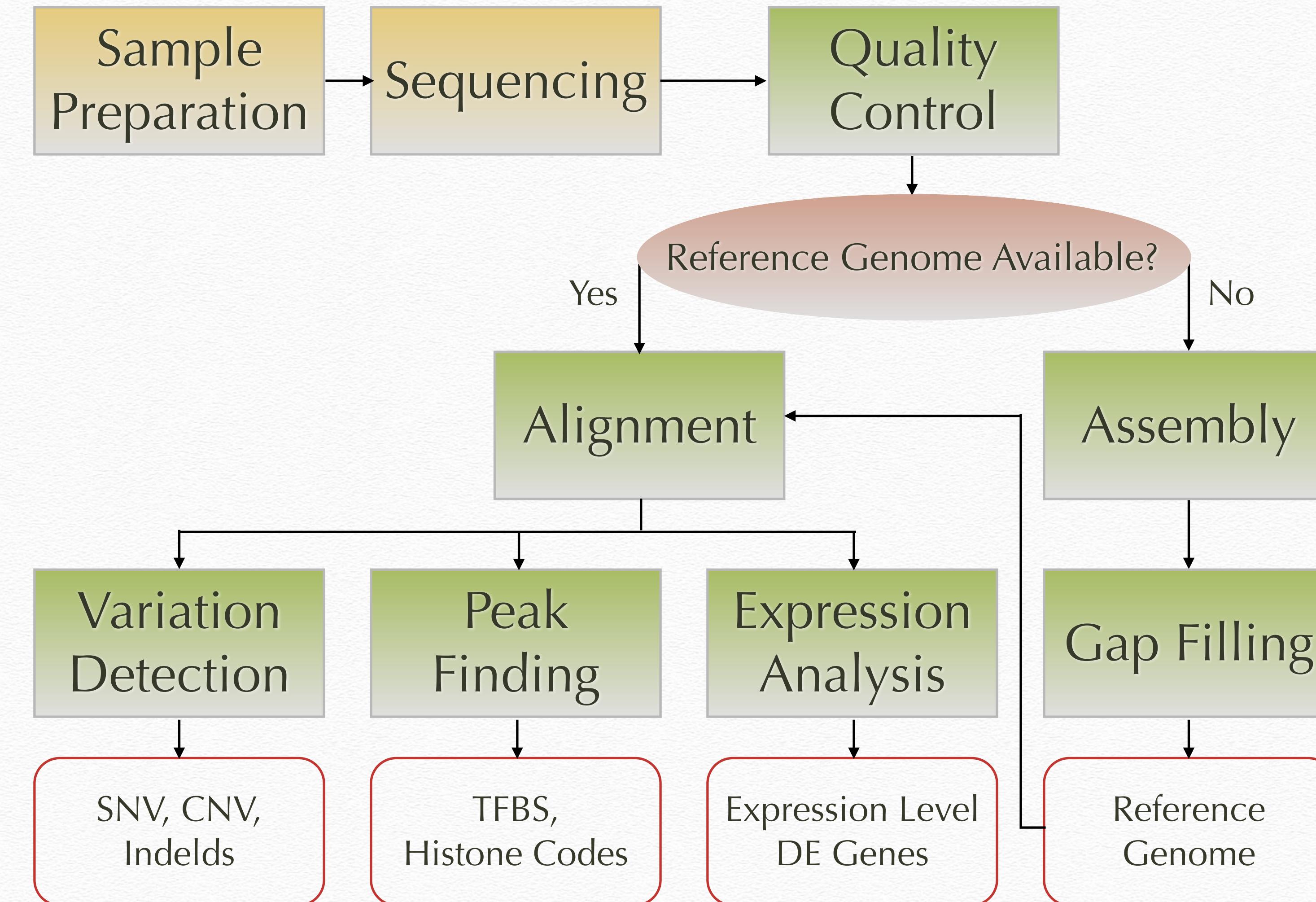
DeepVariant



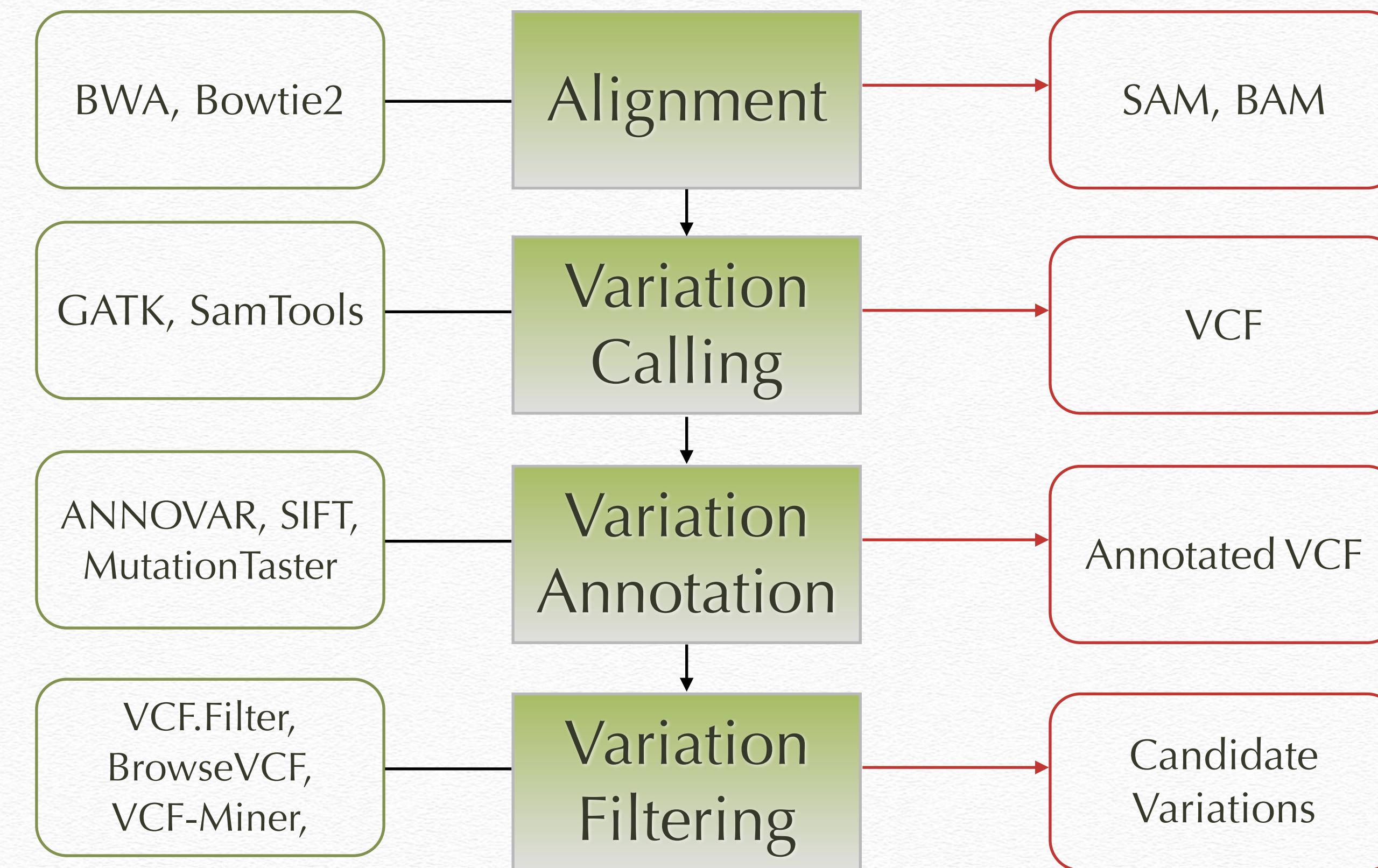
DeepVariant



Analysis Pipeline



Analysis Pipeline



VCF File Structure

#CHROM (required)	POS (required)	ID (optional)	REF (optional)	ALT (optional)	INFO (optional)	Column 7 and beyond (ignored)
Chromosome identifier. Numbers are preferred, but chr or ch prefixes are allowed. All cells in this column must be filled.	Position in the reference sequence. All cells in this column must be filled.	All cells must contain either: <ul style="list-style-type: none"> • For known dbSNP entries, the rs ID • For unknown or nonexistent IDs, a period (.) 	The reference base(s) For unknown bases, a period (.)	The variant base(s) For unknown bases, a period (.)	User ID and source assembly information For unknown bases, a period (.)	These columns may contain data, but they will be ignored by the SeqMan NGen aligner and variant caller.

VCF File Structure

VCF File Structure

	#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	salma
483	chr1	13550	.	G	A	355.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-0.527;ClippingRankSum=0.000;DP=83;ExcessHet=3.0103;FS=8.417;MLEAC=1;MLEAF=0.50		
484	chr1	13813	.	T	G	55.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.660;ClippingRankSum=0.000;DP=11;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500		
485	chr1	13838	.	C	T	144.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=1.550;ClippingRankSum=0.000;DP=11;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500		
486	chr1	14464	.	A	T	286.78	.	AC=2;AF=1.00;AN=2;DP=8;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=36.50;QD=34.24;SOR=2.833	GT:AD:DP:GQ:PL	
487	chr1	14574	.	A	G	38.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-0.725;ClippingRankSum=0.000;DP=31;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.50		
488	chr1	14590	.	G	A	124.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=1.545;ClippingRankSum=0.000;DP=36;ExcessHet=3.0103;FS=5.200;MLEAC=1;MLEAF=0.500		
489	chr1	14599	.	T	A	106.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=1.253;ClippingRankSum=0.000;DP=42;ExcessHet=3.0103;FS=7.988;MLEAC=1;MLEAF=0.500		
490	chr1	14604	.	A	G	91.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=2.702;ClippingRankSum=0.000;DP=48;ExcessHet=3.0103;FS=8.028;MLEAC=1;MLEAF=0.500		
491	chr1	14610	.	T	C	100.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.776;ClippingRankSum=0.000;DP=55;ExcessHet=3.0103;FS=3.917;MLEAC=1;MLEAF=0.500		
492	chr1	14653	.	C	T	259.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=1.952;ClippingRankSum=0.000;DP=89;ExcessHet=3.0103;FS=5.634;MLEAC=1;MLEAF=0.500		
493	chr1	14932	.	G	T	103.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-0.126;ClippingRankSum=0.000;DP=31;ExcessHet=3.0103;FS=8.321;MLEAC=1;MLEAF=0.50		
494	chr1	14937	.	T	C	103.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.693;ClippingRankSum=0.000;DP=31;ExcessHet=3.0103;FS=8.768;MLEAC=1;MLEAF=0.500		
495	chr1	15274	.	A	G	131.77	.	AC=2;AF=1.00;AN=2;DP=6;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500;MQ=25.18;QD=32.94;SOR=3.258	GT:AD:DP:GQ	
496	chr1	15903	.	G	GC	319.73	.	AC=2;AF=1.00;AN=2;DP=10;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=33.28;QD=30.63;SOR=2.093	GT:AD:DP:GQ	
497	chr1	16495	.	G	C	78.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-0.589;ClippingRankSum=0.000;DP=9;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500		
498	chr1	16949	.	A	C	355.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.194;ClippingRankSum=0.000;DP=74;ExcessHet=3.0103;FS=10.432;MLEAC=1;MLEAF=0.50		
499	chr1	17020	.	G	A	197.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-1.193;ClippingRankSum=0.000;DP=69;ExcessHet=3.0103;FS=6.116;MLEAC=1;MLEAF=0.50		
500	chr1	17385	.	G	A	122.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.324;ClippingRankSum=0.000;DP=62;ExcessHet=3.0103;FS=1.495;MLEAC=1;MLEAF=0.500		
501	chr1	17594	.	C	T	198.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=1.810;ClippingRankSum=0.000;DP=115;ExcessHet=3.0103;FS=3.868;MLEAC=1;MLEAF=0.50		
502	chr1	17626	.	G	A	118.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-2.346;ClippingRankSum=0.000;DP=84;ExcessHet=3.0103;FS=3.494;MLEAC=1;MLEAF=0.50		
503	chr1	17697	.	G	C	310.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=-0.259;ClippingRankSum=0.000;DP=58;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.50		
504	chr1	19004	.	A	G	50.28	.	AC=2;AF=1.00;AN=2;DP=3;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=21.35;QD=16.76;SOR=2.833	GT:AD:DP:GQ:PL	
505	chr1	19322	.	C	T	85.77	.	AC=1;AF=0.500;AN=2;BaseQRankSum=0.289;ClippingRankSum=0.000;DP=12;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500		
506	chr1	98921	.	AG	A	322.75	.	AC=2;AF=1.00;AN=2;DP=9;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=49.09;QD=29.09;SOR=1.179	GT:AD:DP:GQ:PL	
507	chr1	98933	.	T	A	466.77	.	AC=2;AF=1.00;AN=2;DP=10;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=52.24;QD=32.93;SOR=0.693	GT:AD:DP:GQ	
508	chr1	98945	.	C	T	466.77	.	AC=2;AF=1.00;AN=2;DP=11;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=52.99;QD=30.83;SOR=0.859	GT:AD:DP:GQ	
509	chr1	127835	.	T	C	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
510	chr1	127837	.	T	C	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
511	chr1	127838	.	G	A	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
512	chr1	127844	.	T	C	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
513	chr1	127845	.	G	A	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
514	chr1	127849	.	A	T	62.74	.	AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=40.00;QD=31.37;SOR=0.693	GT:AD:DP:GQ:PL	
515	chr1	128798	.	C	T	46.28	.	AC=2;AF=1.00;AN=2;DP=3;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=35.03;QD=15.43;SOR=1.179	GT:AD:DP:GQ:PL	
516	chr1	129010	.	AATG	A	322.75	.	AC=2;AF=1.00;AN=2;DP=8;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=20.00;QD=29.67;SOR=2.833	GT:AD:DP:GQ	
517	chr1	129285	.	G	A	540.77	.	AC=2;AF=1.00;AN=2;DP=23;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=22.69;QD=23.51;SOR=0.963	GT:AD:DP:GQ	
518	chr1	133129	.	G	A	135.03	.	AC=2;AF=1.00;AN=2;DP=4;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=43.13;QD=33.76;SOR=3.258	GT:AD:DP:GQ:PL	
519	chr1	138156	.	G	T	193.78	.	AC=2;AF=1.00;AN=2;DP=8;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=23.02;QD=24.22;SOR=2.833	GT:AD:DP:GQ:PL	

SIFT



SIFT

JCVI Home SIFT Home Help Team Contact us

→ SIFT Home
→ Help
→ Contact us

Code release
License
Source Code JCVI-SIFT v. 1.03
Code & exe (Sun, Linux)

FTP download
SIFT Human DB (release 63)
SIFT dbSNP DB (build 132)

Related links
Human genome assembly
GRCh37
Ensembl annotation release 63
NCBI dbSNP Build 132
NCBI BLink

Updates
Aug 2011: SIFT Human DB updated to support GRCh37 Ensembl release 63

Apr 2011: SIFT dbSNP DB updated to support NCBI dbSNP build 132

SIFT predicts whether an amino acid substitution affects protein function. SIFT prediction is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST. SIFT can be applied to naturally occurring **nonsynonymous polymorphisms or laboratory-induced missense mutations**.

**** PROVEAN project New ****

Visit our new [PROVEAN project](#) to get functional predictions from multiple tools. We welcome your [feedback or questions](#).

New features in PROVEAN Human Genome Variants DB:

- Single submission returns functional predictions from **SIFT** and **PROVEAN**. PROVEAN is a new prediction tool which works for **both SNPs and indels**. (Choi et al., 2012, PLOS ONE)
- Updated versions of Ensembl gene annotation (GRCh37 Ensembl 66) and NCBI dbSNP database (Build 137).
- New database structure to support fast retrieval for genome-wide analysis.

Human Genome DB	Tool Description
SIFT/PROVEAN Human SNPs	Get SIFT and PROVEAN predictions for SNPs and indels (Ensembl 66) (Sample format)
SIFT Human SNPs	Get SIFT predictions for nonsynonymous SNPs (Ensembl 63) (Sample format)
	Other human genome tools: <ul style="list-style-type: none">Restrict to Coding Variants (Sample format)Classify Human indels (Sample format)
SIFT Human Protein DB	Tool Description (Ensembl 63)
SIFT Human Protein	Get SIFT predictions for nonsynonymous AA substitutions (Ensembl ENSP ID)
SIFT dbSNP DB	Tool Description (dbSNP Build 132)
SIFT dbSNP rs IDs	Get SIFT predictions for dbSNP SNPs including non-human species (NCBI rs ID)

MutationTaster



mutation t@sting

Gene

HGNC gene symbol, NCBI Gene ID, Ensembl gene ID [show available transcripts](#)

Transcript

Ensembl transcript ID

Position / snippet refers to

coding sequence (ORF)

transcript (cDNA sequence)

gene (genomic sequence)

Alteration

all types by sequence

enter a few bases around your alteration

Format:

ACTGTC[A/*T*] GTGTF A substituted by *T*
ACTGTC[AG/*T*] GTGTF AG substituted by *T*
ACTGTC[ACGT/-] GTGTF ACGT deleted
ACTGTC[-/AA] GTGTF AA inserted

options
 show nucleotide alignment

single base exchange by position

enter position

and new base

insertion or deletion by position

enter positions of

...last wild type base before alteration

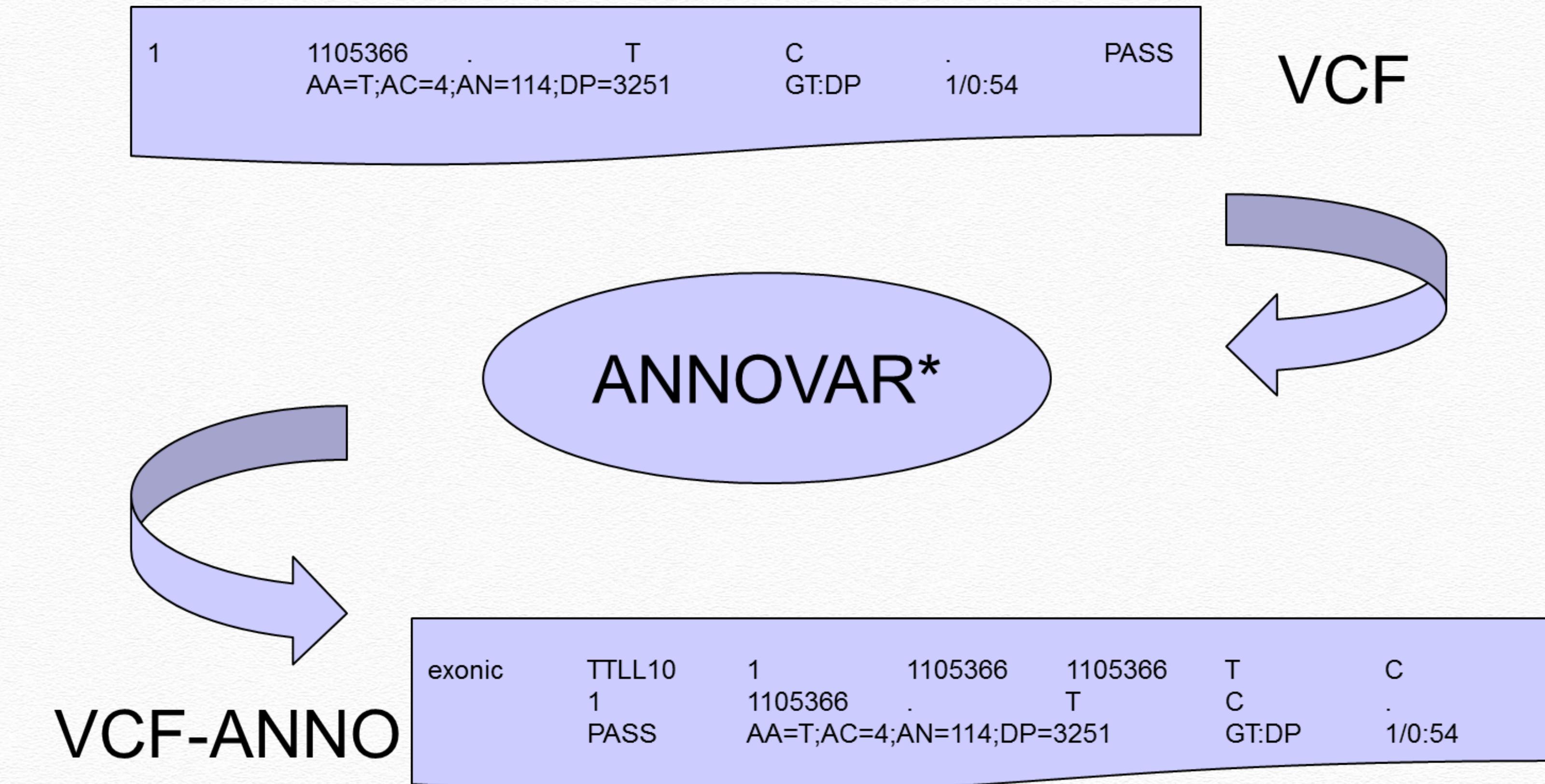
...first wild type base after alteration

and the inserted bases
(if applicable)

Name of alteration

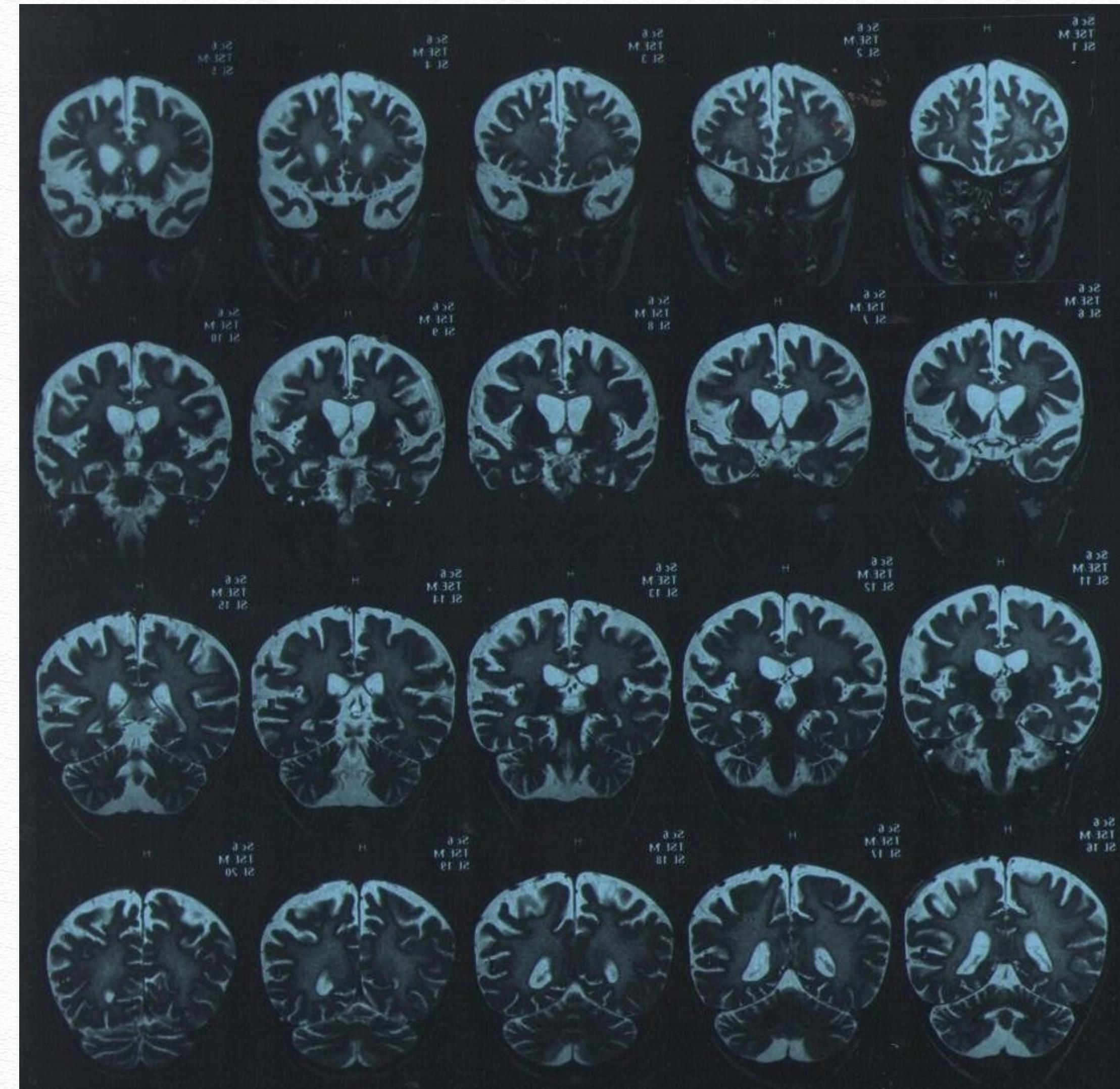
if you would like to have a name for this alteration in the output later on, please type in here

ANNOVAR



*Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data Nucleic Acids Research, 38:e164, 2010
(www.openbioinformatics.org/annovar)

Real Case: Child



Real Case: Child

Report date: 04/12/10																																			
Clinical Features: Clinically diagnosed with metachromatic leukodystrophy (MLD)																																			
Referral Reason: This individual is referred for genetic diagnosis and molecular analysis of <i>ARSA</i> gene for MLD.																																			
Result Summary	<p>[c.190T>A]; [c.190T>A]</p> <p>This patient demonstrated four variants:</p> <table border="1"> <thead> <tr> <th>Mutation</th> <th>Location</th> <th>Rs#</th> <th>Genotype</th> <th>Clinical Significance</th> </tr> <tr> <th>DNA</th> <th>Amino acid</th> <th></th> <th></th> <th></th> </tr> </thead> <tbody> <tr> <td>c.190T>A</td> <td>F64I</td> <td>Exon</td> <td>AA</td> <td>unknown function</td> </tr> <tr> <td>g.6790A>G</td> <td>-</td> <td>Intron</td> <td>Rs762672</td> <td>GG unknown function</td> </tr> <tr> <td>g.7467C>T</td> <td>-</td> <td>Intron</td> <td>Rs762673</td> <td>TT unknown function</td> </tr> <tr> <td>g.7621C>G</td> <td>-</td> <td>Intron</td> <td>Rs762674</td> <td>GG benign</td> </tr> </tbody> </table>					Mutation	Location	Rs#	Genotype	Clinical Significance	DNA	Amino acid				c.190T>A	F64I	Exon	AA	unknown function	g.6790A>G	-	Intron	Rs762672	GG unknown function	g.7467C>T	-	Intron	Rs762673	TT unknown function	g.7621C>G	-	Intron	Rs762674	GG benign
Mutation	Location	Rs#	Genotype	Clinical Significance																															
DNA	Amino acid																																		
c.190T>A	F64I	Exon	AA	unknown function																															
g.6790A>G	-	Intron	Rs762672	GG unknown function																															
g.7467C>T	-	Intron	Rs762673	TT unknown function																															
g.7621C>G	-	Intron	Rs762674	GG benign																															
Interpretation and Comments	<ul style="list-style-type: none"> Direct sequencing of coding region of <i>ARSA</i> gene of this patient revealed four variants including c.190T>A, rs762674, rs762672 and rs762673. c.190T>A variant, F64I (Phenylalanine to Isoleucine) is predicted to be damaging by SIFT, disease causing by MutationTaster and deleterious by PROVEAN software. Most individuals with MLD have mutations in the <i>ARSA</i> gene, which provides instructions for making the enzyme arylsulfatase A which helps break down sulfatides. A few individuals with MLD have mutations in the <i>PSAP</i> gene that provides instructions for making a protein that is broken up (cleaved) into smaller proteins that assist enzymes in breaking down various fats. Mutations in the <i>ARSA</i> or <i>PSAP</i> genes result in a decreased ability to break down sulfatides, resulting in the accumulation of these substances in cells. It is highly recommended that segregation analysis for this variant c.190T>A should be performed for other family members. It should be noted that clinical diagnosis would change the results of the candidate gene for molecular confirmation of the disease. As this result has important implication for the family, a referral to a clinical genetic service is recommended. 																																		
Analysis Method	<ul style="list-style-type: none"> DNA was extracted using salting out method. Direct sequencing of coding region of <i>ARSA</i> gene (exons and exonic/intronic boundaries) was performed, provided by Genefanavar Co. with the sensitivity of 99%. The mutations were named using <i>ARSA</i> gene NCBI Reference Sequence: NG_009260.2 The current description of mutations is as recommended by HGVS. The accuracy of the result depends on the clinical diagnosis and the assumption that the <i>ARSA</i> gene is responsible for MLD. 																																		

Real Case: Father

<p>Clinical Features: No significant clinical symptoms. Family history showed having a daughter with MLD.</p> <p>Referral Reason: This individual is referred for genetic diagnosis and molecular analysis of of <i>ARSA</i> gene due to having a daughter with MLD</p>																																								
<p>Result [c.190T>A];[N] N=Normal</p>																																								
<p>Summary This patient demonstrated five variants:</p> <table border="1"><thead><tr><th>Mutation</th><th>Location</th><th>Rs#</th><th>Genotype</th><th>Clinical Significance</th></tr></thead><tbody><tr><td>DNA</td><td>Amino acid</td><td></td><td></td><td></td></tr><tr><td>c.190T>A</td><td>F64I</td><td>Exon</td><td>TA</td><td>unknown function</td></tr><tr><td>g.7369C>T</td><td>-</td><td>Intron</td><td>Rs6151423 CT</td><td>unknown function</td></tr><tr><td>g.6790A>G</td><td>-</td><td>Intron</td><td>Rs762672 GG</td><td>unknown function</td></tr><tr><td>g.7467C>T</td><td>-</td><td>Intron</td><td>Rs762673 TT</td><td>unknown function</td></tr><tr><td>g.7621C>G</td><td>-</td><td>Intron</td><td>Rs762674 GG</td><td>benign</td></tr></tbody></table>						Mutation	Location	Rs#	Genotype	Clinical Significance	DNA	Amino acid				c.190T>A	F64I	Exon	TA	unknown function	g.7369C>T	-	Intron	Rs6151423 CT	unknown function	g.6790A>G	-	Intron	Rs762672 GG	unknown function	g.7467C>T	-	Intron	Rs762673 TT	unknown function	g.7621C>G	-	Intron	Rs762674 GG	benign
Mutation	Location	Rs#	Genotype	Clinical Significance																																				
DNA	Amino acid																																							
c.190T>A	F64I	Exon	TA	unknown function																																				
g.7369C>T	-	Intron	Rs6151423 CT	unknown function																																				
g.6790A>G	-	Intron	Rs762672 GG	unknown function																																				
g.7467C>T	-	Intron	Rs762673 TT	unknown function																																				
g.7621C>G	-	Intron	Rs762674 GG	benign																																				
<p>Interpretation and Comments</p> <ul style="list-style-type: none">• Direct sequencing of coding region of <i>ARSA</i> gene of this patient revealed five variants including c.190T>A, Rs6151423, Rs762672, Rs762673 and Rs762674.• Most individuals with MLD have mutations in the <i>ARSA</i> gene, which provides instructions for making the enzyme arylsulfatase A which helps break down sulfatides.• A few individuals with MLD have mutations in the <i>PSAP</i> gene that provides instructions for making a protein that is broken up (cleaved) into smaller proteins that assist enzymes in breaking down various fats. Mutations in the <i>ARSA</i> or <i>PSAP</i> genes result in a decreased ability to break down sulfatides, resulting in the accumulation of these substances in cells.• It is highly recommended that segregation analysis for this variant c.190T>A should be performed for other family members.• It should be noted that clinical diagnosis would change the results of the candidate gene for molecular confirmation of the disease.• As this result has important implication for the family, a referral to a clinical genetic service is recommended.																																								
<p>Analysis Method</p> <ul style="list-style-type: none">▪ DNA was extracted using salting out method.▪ Direct sequencing of coding region of <i>ARSA</i> gene (exons and exonic/intronic boundaries) was performed, provided by Genefanavar Co. with the sensitivity of 99%.▪ The mutations were named using <i>ARSA</i> gene NCBI Reference Sequence: NG_009260.2▪ The current description of mutations is as recommended by HGVS.▪ The accuracy of the result depends on the clinical diagnosis and the assumption that the <i>ARSA</i> gene is responsible for MLD.																																								

Real Case: Mother

<p>Clinical Features: No significant clinical symptoms. Family history showed having a daughter with MLD.</p> <p>Referral Reason: This individual is referred for genetic diagnosis and molecular analysis of of <i>ARSA</i> gene due to having a daughter with MLD</p>																																																		
Result	[c.190T>A];[N] N=Normal																																																	
Summary	<p>This patient demonstrated six variants:</p> <table border="1"> <thead> <tr> <th>Mutation</th> <th>Location</th> <th>Rs#</th> <th>Genotype</th> <th>Clinical Significance</th> </tr> <tr> <th>DNA</th> <th>Amino acid</th> <th></th> <th></th> <th></th> </tr> </thead> <tbody> <tr> <td>c.190T>A</td> <td>F64I</td> <td>Exon</td> <td>TA</td> <td>unknown function</td> </tr> <tr> <td>g.6253C>T</td> <td>-</td> <td>Intron</td> <td>Rs34457249</td> <td>CT</td> <td>unknown function</td> </tr> <tr> <td>g.7569C>G</td> <td>Thr393Ser</td> <td>Exon</td> <td>Rs743616</td> <td>CG</td> <td>Conflicting interpretations of pathogenicity (1,2).</td> </tr> <tr> <td>g.7621C>G</td> <td>-</td> <td>Intron</td> <td>Rs762674</td> <td>GG</td> <td>benign</td> </tr> <tr> <td>g.6790A>G</td> <td>-</td> <td>Intron</td> <td>Rs762672</td> <td>GG</td> <td>unknown function</td> </tr> <tr> <td>g.7467C>T</td> <td>-</td> <td>Intron</td> <td>Rs762673</td> <td>TT</td> <td>unknown function</td> </tr> </tbody> </table>					Mutation	Location	Rs#	Genotype	Clinical Significance	DNA	Amino acid				c.190T>A	F64I	Exon	TA	unknown function	g.6253C>T	-	Intron	Rs34457249	CT	unknown function	g.7569C>G	Thr393Ser	Exon	Rs743616	CG	Conflicting interpretations of pathogenicity (1,2).	g.7621C>G	-	Intron	Rs762674	GG	benign	g.6790A>G	-	Intron	Rs762672	GG	unknown function	g.7467C>T	-	Intron	Rs762673	TT	unknown function
Mutation	Location	Rs#	Genotype	Clinical Significance																																														
DNA	Amino acid																																																	
c.190T>A	F64I	Exon	TA	unknown function																																														
g.6253C>T	-	Intron	Rs34457249	CT	unknown function																																													
g.7569C>G	Thr393Ser	Exon	Rs743616	CG	Conflicting interpretations of pathogenicity (1,2).																																													
g.7621C>G	-	Intron	Rs762674	GG	benign																																													
g.6790A>G	-	Intron	Rs762672	GG	unknown function																																													
g.7467C>T	-	Intron	Rs762673	TT	unknown function																																													
Interpretation and Comments	<ul style="list-style-type: none"> Direct sequencing of coding region of <i>ARSA</i> gene of this patient revealed six variants including c.190T>A, Rs34457249, Rs743616, Rs762674, Rs762672 and Rs762673. Most individuals with MLD have mutations in the <i>ARSA</i> gene, which provides instructions for making the enzyme arylsulfatase A which helps break down sulfatides. A few individuals with MLD have mutations in the <i>PSAP</i> gene that provides instructions for making a protein that is broken up (cleaved) into smaller proteins that assist enzymes in breaking down various fats. Mutations in the <i>ARSA</i> or <i>PSAP</i> genes result in a decreased ability to break down sulfatides, resulting in the accumulation of these substances in cells. It is highly recommended that segregation analysis for this variant c.190T>A should be performed for other family members. It should be noted that clinical diagnosis would change the results of the candidate gene for molecular confirmation of the disease. As this result has important implication for the family, a referral to a clinical genetic service is recommended. <p>1) http://www.ncbi.nlm.nih.gov/clinvar/variation/21184/#clinical-assertions 2) Bean LJ, et al, Hum Mutat. 2013 Sep;34(9):1183-8.</p>																																																	
Analysis Method	<ul style="list-style-type: none"> DNA was extracted using salting out method. Direct sequencing of coding region of <i>ARSA</i> gene (exons and exonic/intronic boundaries) was performed, provided by Genefanavar Co. with the sensitivity of 99%. The mutations were named using <i>ARSA</i> gene NCBI Reference Sequence: NG_009260.2 The current description of mutations is as recommended by HGVS. 																																																	