



北海道大学

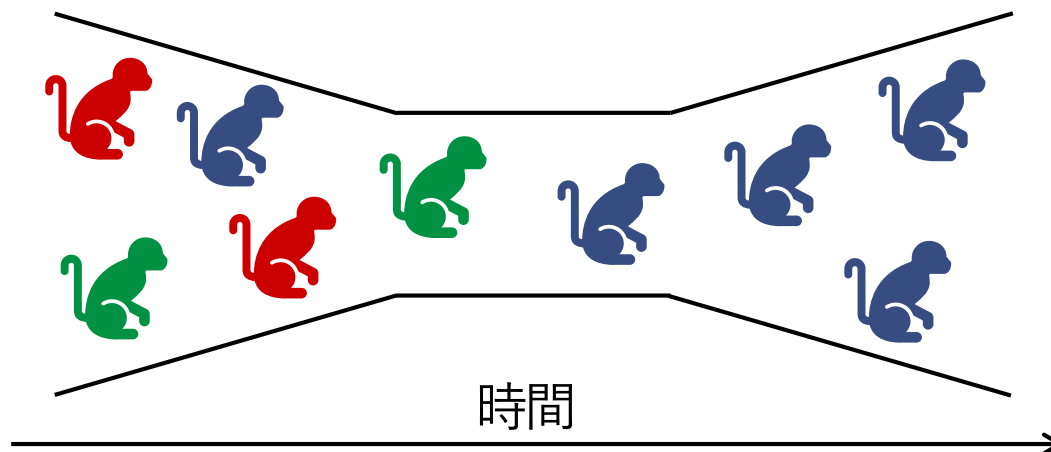
Evolutionary Probability の応用による ヒト適応的アミノ酸変異の検出

目的

- ✓ ヒト適応的アミノ酸変異を高精度に検出する

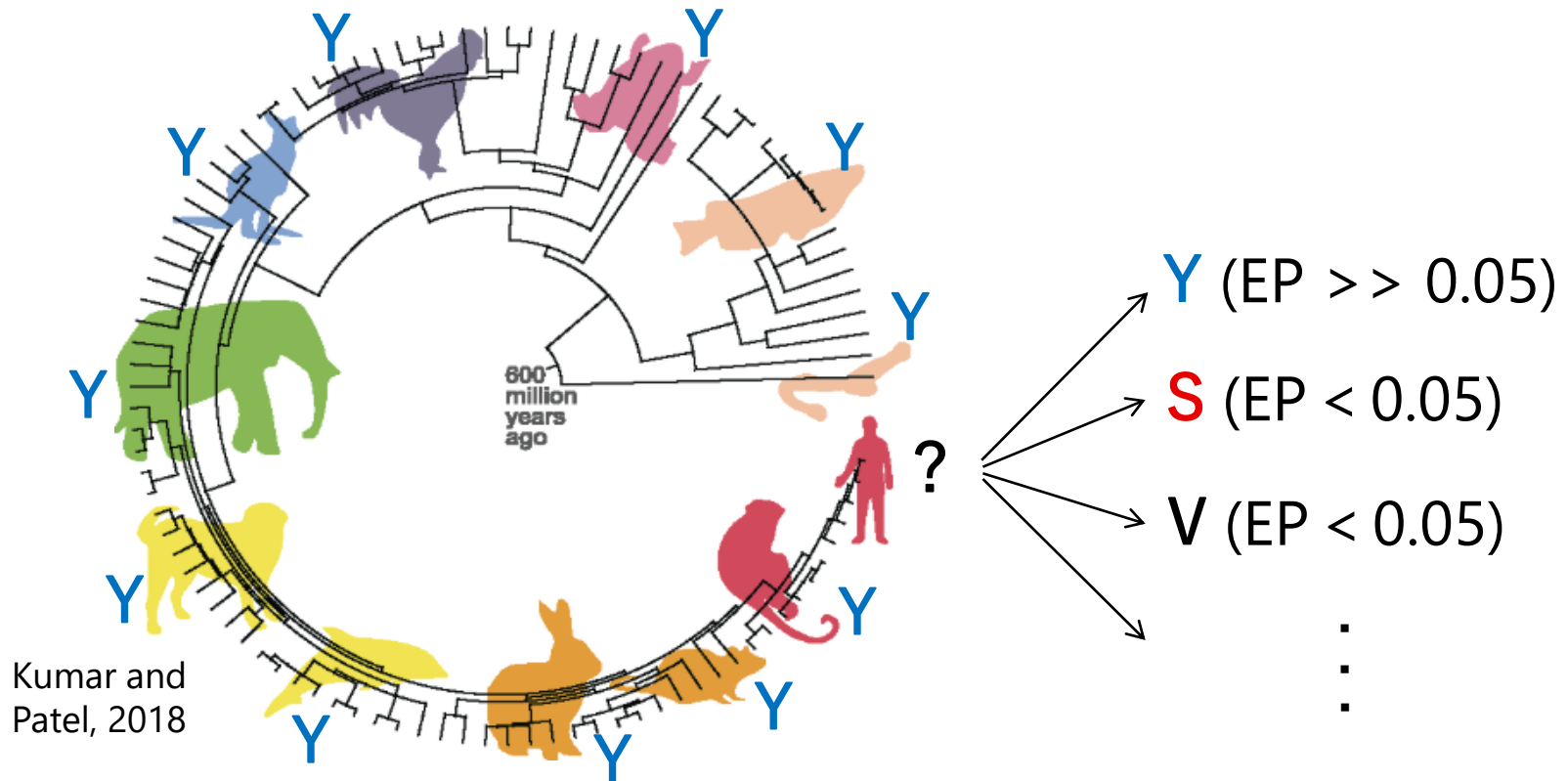
性質や特徴が有利になるようなアミノ酸の個体差

- ✓ 何故そのような進化を遂げたのか解明する手がかりとなる
- ✓ 既存の方法はボトルネック効果の影響を受けてしまう



Evolutionary Probability (EP)

- ✓ あるタンパク質の1アミノ酸を未知だと仮定した際
どのアミノ酸となるのが尤もらしいかを表した確率 (1)
- ✓ $EP < 0.05$ のとき、**有害または適応的な**アミノ酸置換

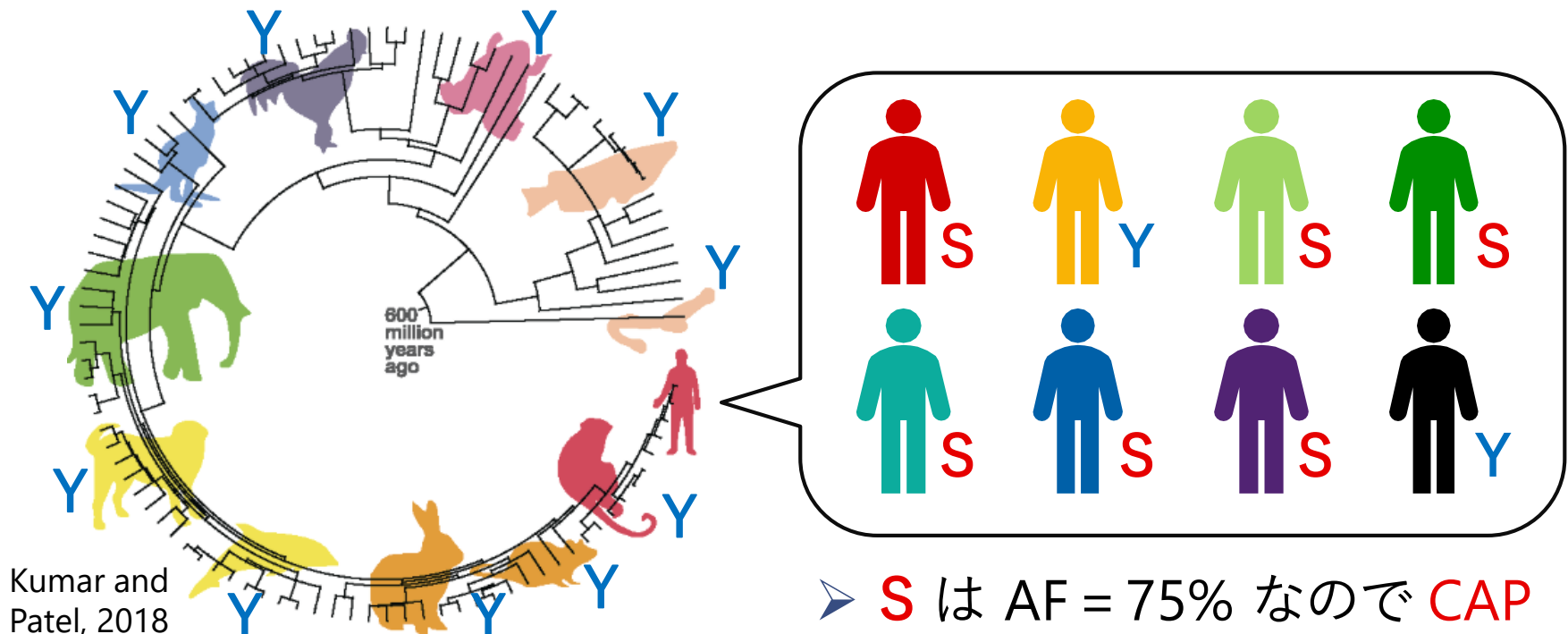


(1) Liu et al. 2016. *Molecular Biology and Evolution*.

Candidate Adaptive Polymorphism (CAP)

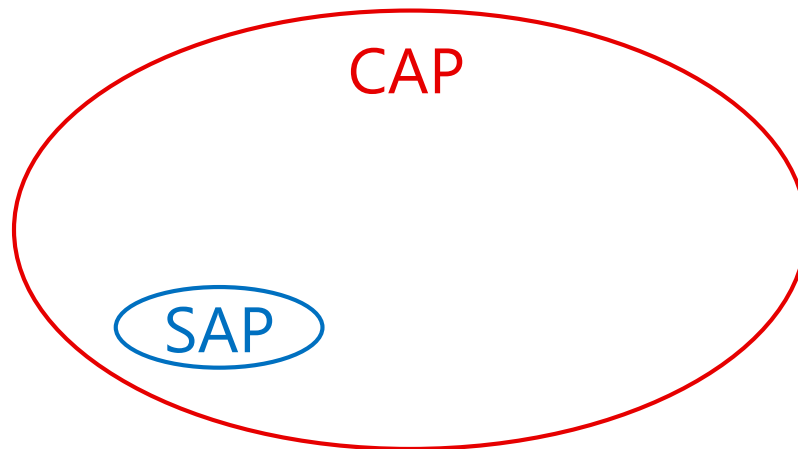
- ✓ 有害または適応的で、集団内で高い頻度のアミノ酸変異 (2)
 $EP < 0.05$ アレル頻度 $AF > 5\%$

- ✓ CAP は適応的変異候補



CAP の問題点

- ✓ CAP に対して、先行研究で示された適応的変異を SAP (Suggested Adaptive Polymorphism) とする
例 | *LOXL1* にある SAP は、緑内障リスクを 20 倍低下
- ✓ CAP の個数は SAP の個数の約 200 倍もあり
適応的でないと考えられる CAP が多く含まれてしまっている



材料と本研究の流れ

材料

- ✓ 先行研究で示された **CAP** データ

mypcg <http://www.mypcg.info/caps> より取得

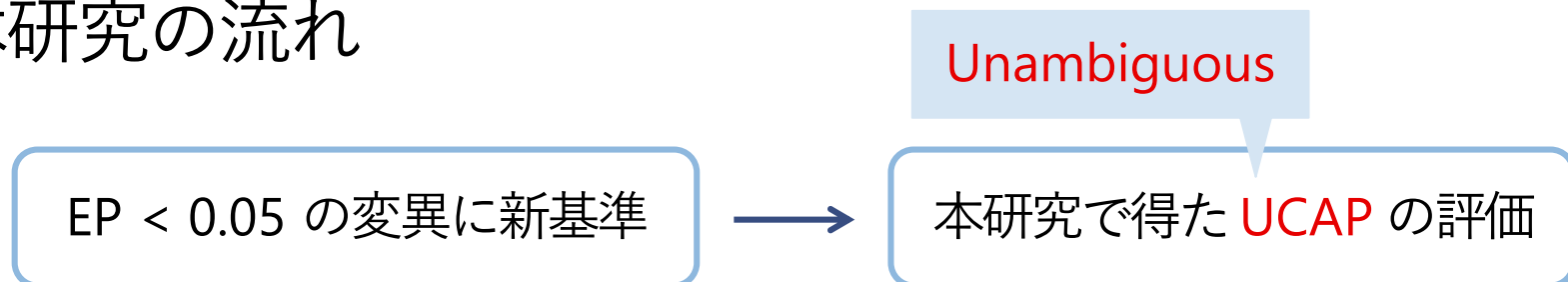
- ✓ 遺伝子配列データ

NCBI <https://www.ncbi.nlm.nih.gov/> より取得

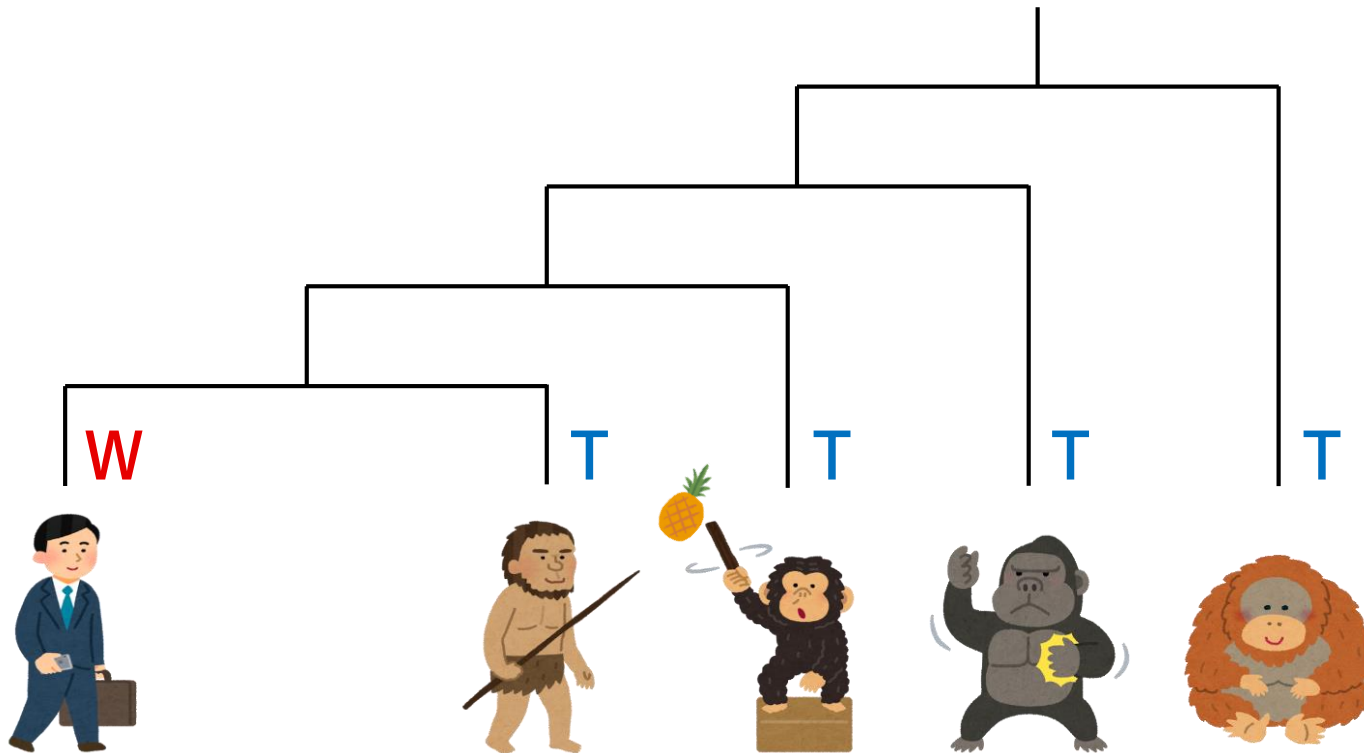
- ✓ 世界 5 地域 2,504 人の全ゲノムデータ

1000 Genomes Project <https://www.internationalgenome.org/> より取得

本研究の流れ

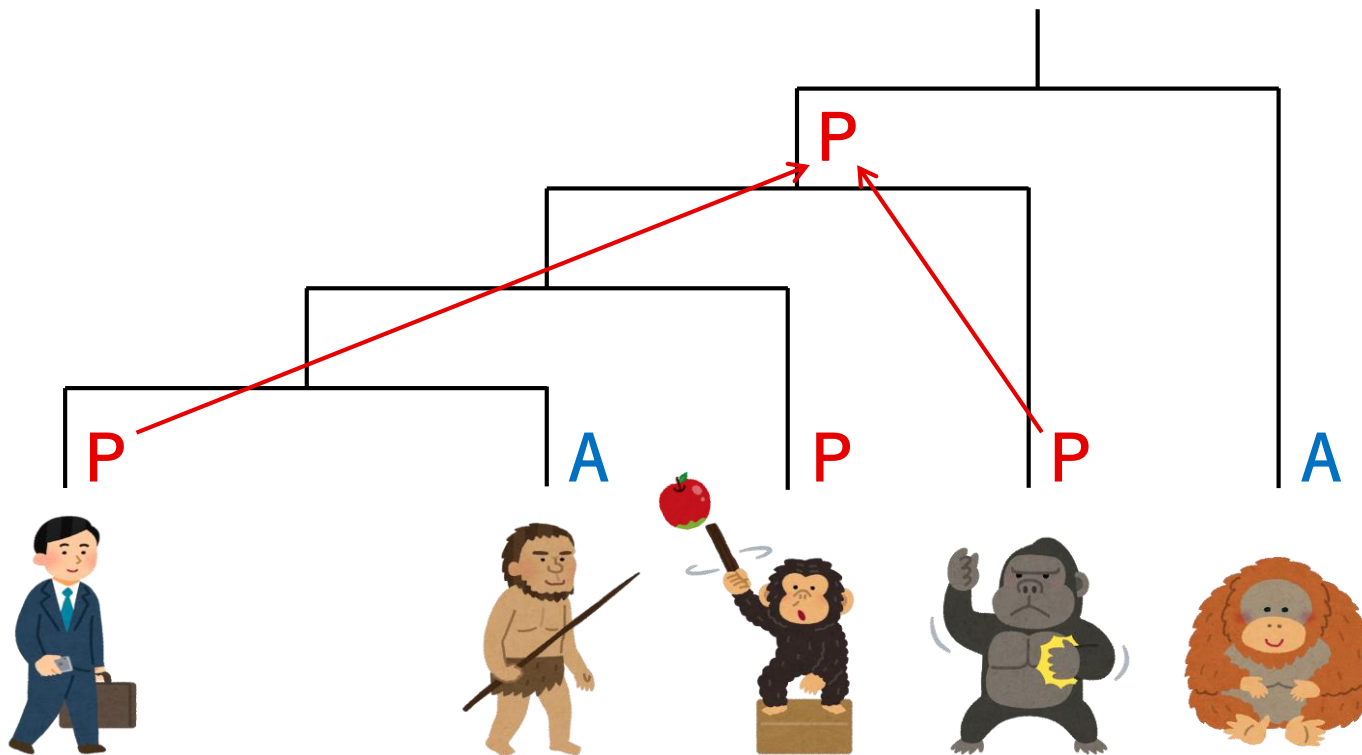


- ✓ 現生人と近縁種の共通祖先で生じた変異を対象外とした
 - 現生人特有の適応的変異を発見するため



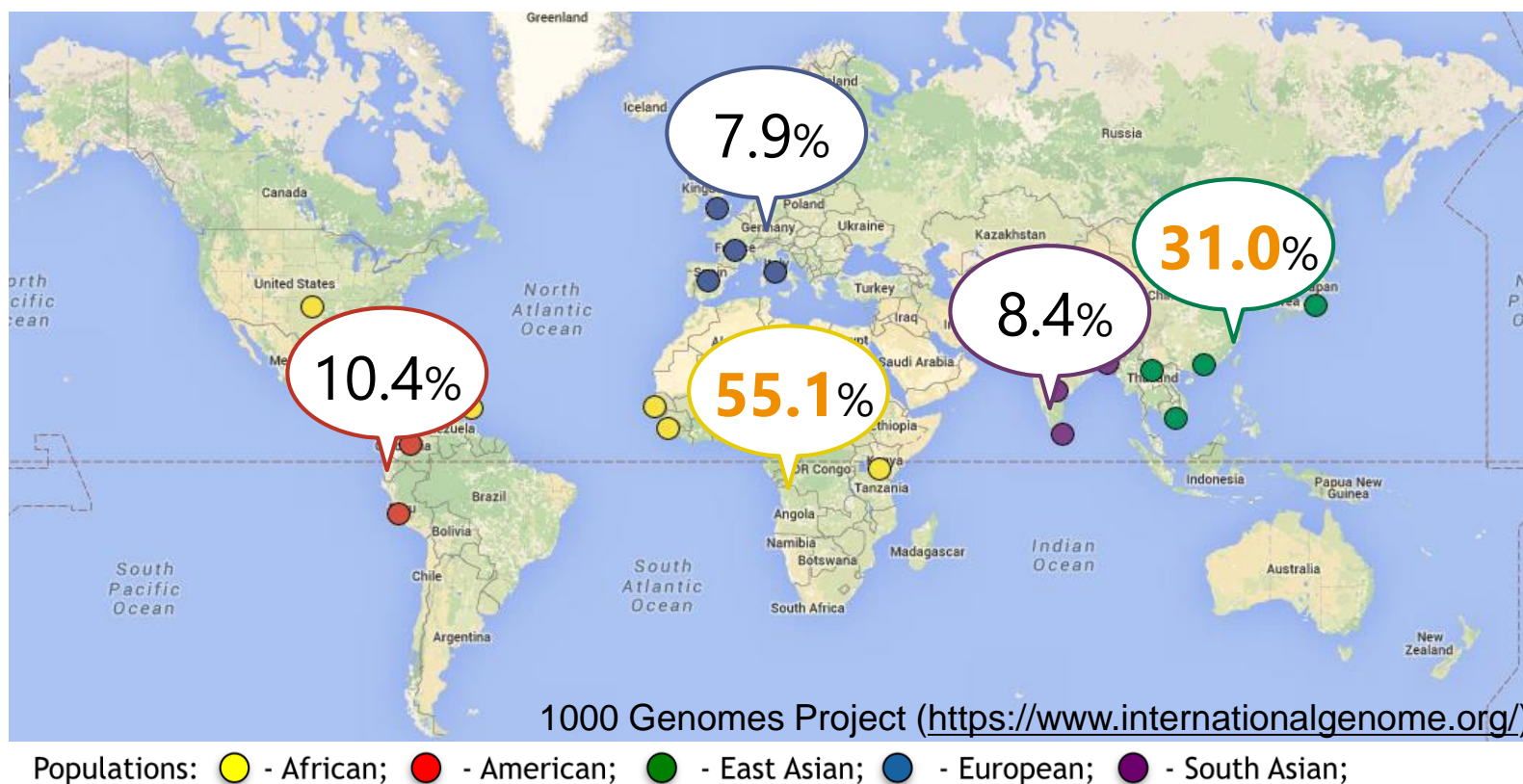
- ✓ 現生人で生じた変異なため**扱う対象**

- ✓ 現生人と近縁種の共通祖先で生じた変異を対象外とした
 - 現生人特有の適応的変異を発見するため



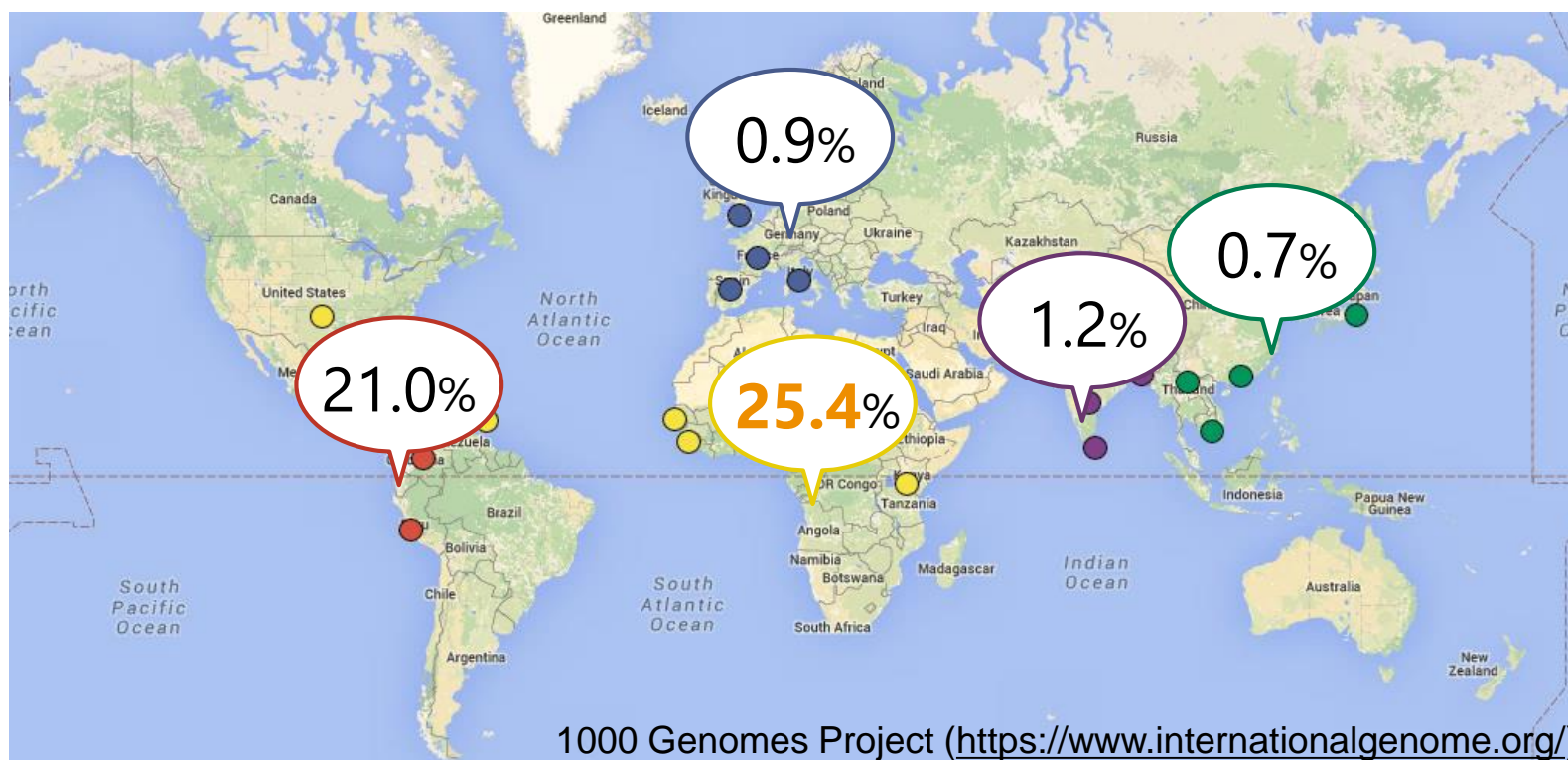
- ✓ 現生人とゴリラの共通祖先で生じた変異なため**対象外**

- ✓ 複数集団でアレル頻度 $AF > 25\%$ となる変異を対象とした
 - 単一集団のみで偶然広まった変異を除去するため



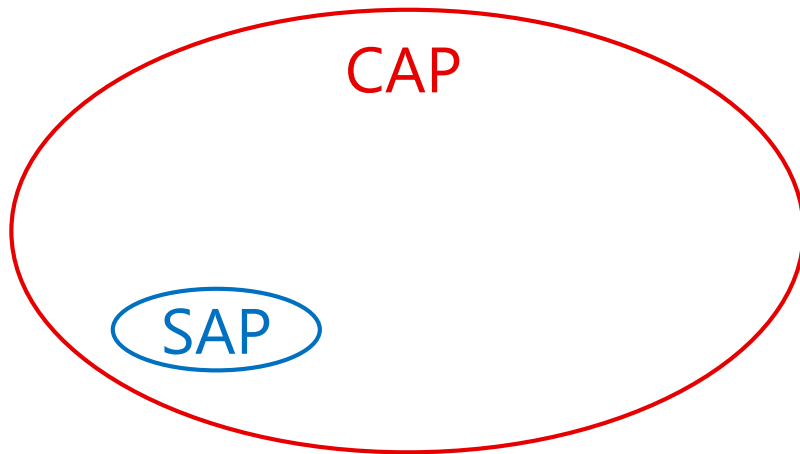
- ✓ アフリカ・東アジア集団で $AF > 25\%$ のため扱う対象

- ✓ 複数集団でアレル頻度 $AF > 25\%$ となる変異を対象とした
 - 単一集団のみで偶然広まった変異を除去するため



Populations: ● - African; ● - American; ● - East Asian; ● - European; ● - South Asian;

- ✓ アフリカ集団のみで $AF > 25\%$ のため対象外



SAP	CAP	SAP / CAP
124	18,724	0.00662

① 近縁種の共通祖先変異を除去

のみ

95

10,101

0.00941

② 複数集団で AF > 25% を残す

のみ

82

12,188

0.00672

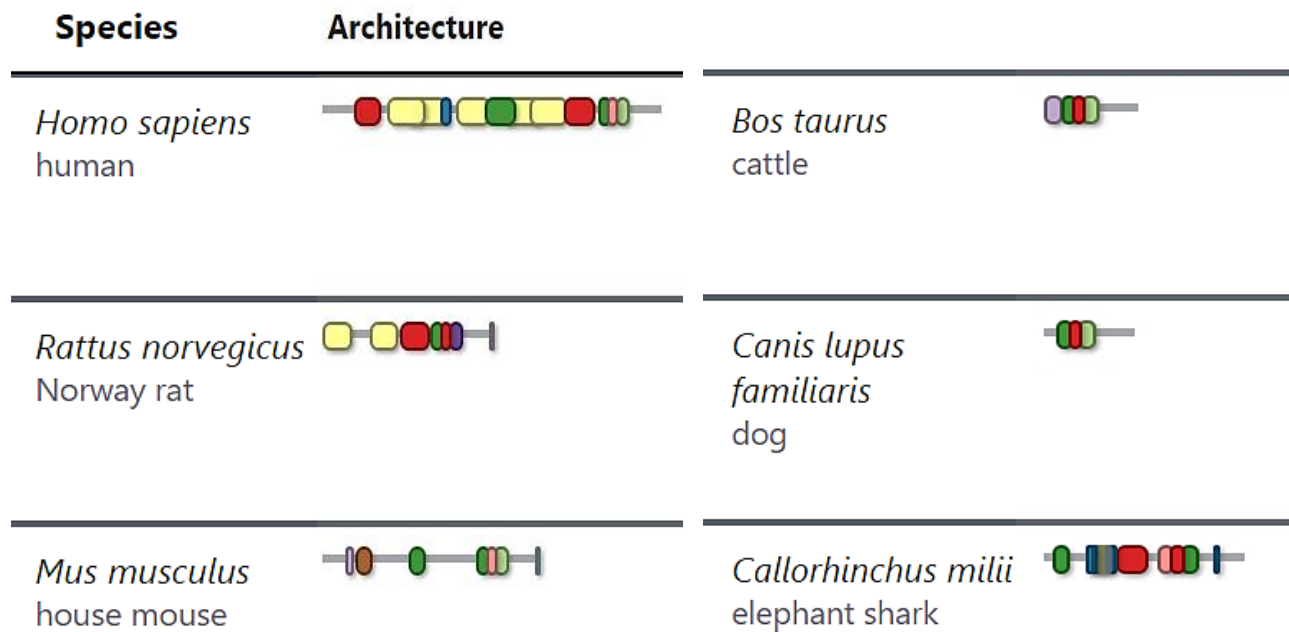


UCAP	SAP / UCAP
63	4,800
	0.01312

約 **2** 倍

- ✓ 除去した変異の大部分は
種ごとにドメイン構造が大きく異なるアミノ酸に含まれていた

例 | *MUC4*



NCBI
(<https://www.ncbi.nlm.nih.gov/>)

- ✓ 構造の保存が必要でないため、中立と思われる変異を多く含む



中立と思われる変異を多く除去

先行研究				本研究			
遺伝子	CAP 数	SAP 数	種間の 構造保存性	遺伝子	UCAP 数	SAP 数	種間の 構造保存性
<i>MUC4</i>	261	0	×	<i>MUC12</i>	17	0	×
<i>AHNAK2</i>	153	0	×	<i>AHNAK2</i>	16	0	×
<i>FLG</i>	90	0	×	<i>MUC4</i>	15	0	×
<i>MUC17</i>	74	0	×	<i>PKD1L2</i>	14	0	○
<i>MUC12</i>	66	0	×	<i>OBSCN</i>	11	0	○
<i>MUC5B</i>	49	0	○	<i>ALMS1</i>	9	7	○
<i>PKD1L2</i>	49	0	×	<i>ADGRV1</i>	9	0	○
<i>HLA-A</i>	39	0	-	<i>ZAN</i>	9	0	○
<i>FCGBP</i>	34	0	×	<i>FCGBP</i>	9	0	×
<i>IGFN1</i>	29	0	×	<i>PCNT</i>	8	0	○
<i>HLA-C</i>	29	0	-	<i>USH2A</i>	8	0	○
<i>HLA-B</i>	28	1	-	<i>FAT1</i>	8	0	○

(遺伝子間のアライメントスコアの正負により '○', '×' を判断、 '-' は種が非常に少ないことを意味)

- ✓ 種間で構造が保存されていない遺伝子が含む変異を除去

先行研究			
遺伝子	CAP 数	SAP 数	種間の 構造保存性
<i>MUC4</i>	261	0	×
<i>AHNAK2</i>	153	0	×
<i>FLG</i>	90	0	×
<i>MUC17</i>	74	0	×
<i>MUC12</i>	66	0	×
<i>MUC5B</i>	49	0	○
<i>PKD1L2</i>	49	0	×
<i>HLA-A</i>	39	0	-
<i>FCGBP</i>	34	0	×
<i>IGFN1</i>	29	0	×
<i>HLA-C</i>	29	0	-
<i>HLA-B</i>	28	1	-

本研究			
遺伝子	UCAP 数	SAP 数	種間の 構造保存性
<i>MUC12</i>	17	0	×
<i>AHNAK2</i>	16	0	×
<i>MUC4</i>	15	0	×
<i>PKD1L2</i>	14	0	○
<i>OBSCN</i>	11	0	○
<i>ALMS1</i>	9	7	○
<i>ADGRV1</i>	9	0	○
<i>ZAN</i>	9	0	○
<i>FCGBP</i>	9	0	×
<i>PCNT</i>	8	0	○
<i>USH2A</i>	8	0	○
<i>FAT1</i>	8	0	○

(遺伝子間のアライメントスコアの正負により '○', '×' を判断、 '-' は種が非常に少ないことを意味)

✓ 正の自然選択を受けている遺伝子だが SAP が見つかっていない



先行研究で示されていない適応的変異を発見した可能性

結論と今後の展望

- ✓ EP < 0.05 の変異に新基準を設定することで
中立と思われる変異を多く除去し
先行研究で示されていない適応的変異を発見した可能性がある
- ✓ ②複数集団で AF > 25% の変異を対象とする基準に関して
系統樹を作成して詳細に集団を区分することで
より高精度な適応的変異の発見が見込める
- ✓ 本研究の手法はヒト以外の種にも応用可能であるため
哺乳類や鳥類をはじめとした脊椎動物でも検証していく