

生命情報科学専門実習 I

～病気に関する遺伝子の探索～

長 浜 バ イ オ 大 学
バ イ オ サ イ エ ン ス 学 部
生 命 情 報 科 学 コ ー ス

バイオ分野の産業や研究分野では医学や医療に係わる課題が益々重要となってきた。世界的な状況としてはゲノム塩基配列を高速で安価に解読する、「次世代シーケンサー」と呼ばれる塩基配列の解読装置が開発されたことで、バイオや医療分野に大きな変化が起きつつある。日米欧の公的研究機関を中心に、2000年に最初のヒトゲノム配列が解読された際には、約10年の年月と世界全体で数千億円の経費がかかったと言われている。ところが、最近登場した次世代シーケンサーを用いれば数百万円の価格と一週間程度の期間で、一個人のゲノム配列が解読可能となっている。さらに、数年以内には、数十万円で一人の個人ゲノムが一日以内で解読可能と予想されている。人間ドックと似た発想で、予防医学の観点から個人のゲノム配列が解読される時期が近いと考えられている。

ヒトゲノム配列の解読、特に特定の個人のゲノム配列の解読は、従来はプライバシーの問題を中心に論じられる傾向にあり、何か怖いことをする研究と考えられがちであった。しかしながら、現在の医学やゲノム科学の発展に伴い、ほとんどの人のゲノム配列内には、なんらかの重篤な病気と係わる遺伝因子を持っていることが明らかとなった。ヒトのゲノム配列が、進化の産物として出来あがったことを考えると、人類の進化の過去の過程において、80歳や100歳まで生きるように淘汰されて来たわけではない。ほとんどの人が100歳を超える程には長生きはしておらず、高齢になればなんらかの病気を発症する。ウイルスや微生物の感染や生活環境が引き起こす病気も多いが、ガンを代表例として、各個人のゲノム配列に起因して発症をする病気も多い。これらの病気について、それを遺伝子やゲノム配列上の多様性の一環として、理解するようになってきている。「ゲノム配列の多様性」が人類集団にとって重要との認識も深まっている。

ゲノム科学や医学の現状を考えると、これからの高齢化社会における予防治療・予防医学においては、希望する個人が、各自のゲノム配列を解読することが有意義と考えられる。さらには、ガンを発症した患者を例にとっても、その患者に適した抗がん剤を処方するための「テーラーメイド・オーダーメイド医療」を行うには、その患者の遺伝子やゲノム配列の特徴を知っておくことが重要となる。プライバシーの問題に十分に配慮すれば、個人ゲノムの解読は、医学・医療に大きく貢献できる。

日米欧の公的研究機関とは別個に、ベンチャー企業としてヒトゲノム解析を独自に行ったベンター博士が、彼のゲノム配列を全世界に公開している。DNAのダブルヘリックス構造をクリック博士とともに発見した、ワトソン博士も彼の配列を公開しているが、彼らの行動は上述の世界の流れを先導しようとの試みと思える。彼らのゲノム配列中にも当然のこととして、疾患の原因となる遺伝子配列の特徴が既に見つかっている。それにも係わらず、あえて公開したことは、予防治療・予防医学の発展の視点からは意義深いことである。しかしながら、ゲノム配列は彼らの親族とも共通性が高いので、全世界に公開することに親族からの異議が提起されているとも聞く。ベンター博士やワトソン博士のように、個人名入りでゲノム配列を世界に公開することの是非には問題があろうが、この問題を十分に承知の上で、彼らは敢えて個人ゲノム情報を公開している。その意味は、ゲノム配列に起因して発症をする病気は、遺伝子やゲノム配列上の多様性の一環として理解すべきであることを伝えるための、先導的な行動とも考えられる。全地球レベルでの生物の多様性が重要なように、ヒトのゲノム配列の多様性も重要と考えられるようになってきている。

今回の本実習では、これらの世界的な動向ならびに、更には1~2年内には次世代シーケンサーを用いて1000人レベルの個人のゲノムが解読される状況を考え、またその解読自体が医療に関係した新しいバイオ産業を創出することを想定して、ヒトゲノム中の病気と関係した遺伝子配列を収集する。次世代シーケンサーは、我が国でも

百台近くが導入済みか導入予定であるが、その次世代シーケンサーが産出する大量なゲノム配列の情報解析をする人材が極端に不足している。この我が国の状況を考え、生命情報の3回生の後期の実習では、次世代シーケンサーが産出する大量配列の情報解析を行っている。

世界の状況には遅れるとは思いますが、我が国でも数年以内には個人のゲノム配列の解読が普及を始めると思える。個人情報管理に十分な配慮を行った上で、個人ゲノム配列の情報解析を行い、「各個人がどのような病気に罹る可能性が高いのか、それをできるだけ避けるにはどのような注意が必要なのか」を、個人別に明らかにすることが重要になる。本実習は、その種のゲノム情報解析の準備と言える。可能であれば、ベンター博士やワトソン博士がどのような病気になる可能性が高いのかを、本実習中に探索してみたい。この種のテーマに驚きを持つかもしれないが、これからの医療を考える上では世界的に重要となる課題である。我が国においては、研究と教育の両面において、世界に遅れをとっている傾向にあるので、敢えて挑戦を試みよう。

病気に関係する遺伝子と言った場合、いかにも「病気を引き起こすための遺伝子」がヒトゲノム中に存在するように聞こえるが、それは誤解である。例えばガンを例にとると、ガンを引き起こす（あるいはガンに関係する）遺伝子と表現した場合、その着目する遺伝子自体が「ガンを引き起こすことを目的として存在している遺伝子」ではない。着目遺伝子自体は、通常はヒトにとって重要な働きをしている遺伝子である。その遺伝子配列の一部に突然変異が起こることで、その遺伝子の生産するタンパク質の機能が変化や欠損を起こし、その結果としてガンを生じさせる場合が考えられる。一方で、タンパク質の配列には一切変化を与えないが、その遺伝子の発現を制御している領域に変化（突然変異）を起こすことで、遺伝子の発現状態が変化し、その結果としてガンを生じさせる例も多い。本来は、発生の特定の時期に特定の臓器でのみ発現をすべき遺伝子が、発生の異なった時期、ないしは異なった臓器で発現してしまうことがガンにつながる例も多い。このような予備的な知識を頭に入れた上で疾患（病気）に係わる遺伝子の探索を行おう。

I. まず、各自で自分の関心のある病気を考えてほしい。ここでは個々人のゲノムの配列に関係する疾患を考えているので、ウイルスや細菌が引き起こす感染症、（例えばインフルエンザや食中毒等）は対象から外す必要がある。各自で関心のある病気について、正確な知識を得ておくためにメルクマニュアル医学百科（Googleで検索せよ。<http://mmh.banyu.co.jp/mmhe2j/index.html>）を参照しておこう。勿論、メルクマニュアル医学百科を見てから、関心のある病気を決めて良い。ガンのように広い範囲の疾患を選択すると以降の操作が難しくなるので、肺ガンや白血病のように、より限定した疾患名を選択する。

Googleで“メルクマニュアル医学百科”をキーワードにして検索する。
次の画面で、メルクマニュアル医学百科[家庭版]をクリックする。

Google

メルクマニュアル医学百科 検索

約 19,800 件 (0.21 秒) 検索オプション

他のキーワード: [メルクマニュアル医学百科](#) [最新家庭版](#)

メルクマニュアル医学百科[家庭版] ← **クリックする**

一般向け医学書。項目別、五十音別、キーワード検索が可能。メーカー提供サイト。
mmh.banyu.co.jp/ - キャッシュ - 類似ページ

消化器の病気 心の健康問題
 皮膚の病気 腎臓と尿路の病気
 女性の健康上の問題 関節
 心臓と血管の病気 知っておきたい基礎知識

[banyu.co.jpからの検索結果](#) »

[メルクマニュアル 家庭版, セクション 1, 知っておきたい基礎知識](#)

お問い合わせ · Worldwide · 万有ホーム · 万有について · 製品情報 · ニュースルーム · 採用情報 · 研究開発情報 · Licensing · [メルクマニュアル家庭版](#) · 米国メルク社とメルクマニュアル · [メルクマニュアル家庭版を検索](#) · [検索](#) · [索引](#) ...

メルクマニュアル医学百科 **最新家庭版**
 THE MERCK MANUAL OF MEDICAL INFORMATION - SECOND HOME EDITION, ONLINE VERSION

メルクマニュアルとは
 医学情報提供の歴史 >>

メルクマニュアル家庭版を検索

白血病

索引

あいうえおかきくけこ

Merck & Co., Inc., Whitehouse Station, N.J., U.S.A.とメルクマニュアル

メルクマニュアルは、医薬品の研究開発力で高い評価を得ているMerck & Co., Inc., Whitehouse Station, N.J., U.S.A.が非営利事業として提供しているもので、世界で最も信頼されている医学書の一つです。

メルクマニュアル医学百科 トップ画面 (<http://mmh.banyu.co.jp/mmhe2j/index.html>)

興味のある病気、例えば“白血病”とキーワードを入力して、検索ボタンをクリックすると、次の検索結果画面が表示される。

最新メルクマニュアル医学百科 **家庭版**

メルクマニュアルとは
 医学情報提供の歴史 >>

メルクマニュアル家庭版を検索

索引

検索結果
 “白血病”

表題からの検索

白血病 ← **例えば“白血病”をクリックする**

白血病は、白血球または成熟して白血球になる細胞のがんです...

急性リンパ球性白血病
 急性リンパ球性白血病は、正常ならリンパ球になる細胞ががん化して、急速に骨髄の正常細胞に取って代わる、命にかかわる病気です...

急性骨髄性白血病
 急性骨髄性(骨髄芽球性、骨髄単球性)白血病は、正常な状態では好中球、好塩基球、好酸球、単球に成長する細胞が、急速に骨髄の正常細胞に取って代わり、命にかかわる

“白血病”をクリックすると、次の解説画面が表示される。

<p>最新 メルクマニュアル 医学百科 家庭版</p>	セクション	血液の病気
	章	白血病
	トピック	はじめに・急性リンパ球性白血病・急性骨髄性白血病・慢性リンパ球性白血病・慢性骨髄性白血病
	検索語	検索語はハイライトしています 白血病

<p>メルクマニュアルとは 医学情報提供の歴史 >></p>	<h2>はじめに</h2> <p>白血病は、白血球または成熟して白血球になる細胞のがんです。</p> <p>白血球は骨髄の幹細胞から生じた細胞です。細胞の成長過程で誤りが生じると、染色体の遺伝子の配列が部分的に変わってしまうことがあります。こうしてできた異常な染色体は、細胞分裂の制御を狂わせ、異常な細胞が無制限に増殖して悪性化し、白血病になります。白血病細胞は最終的に骨髄に侵入して、血球が正常に成長する働きを変えたり抑えたりします。また肝臓、脾臓(ひそう)、リンパ節、精巣、脳などの臓器に侵入することもあります。</p>
--	--

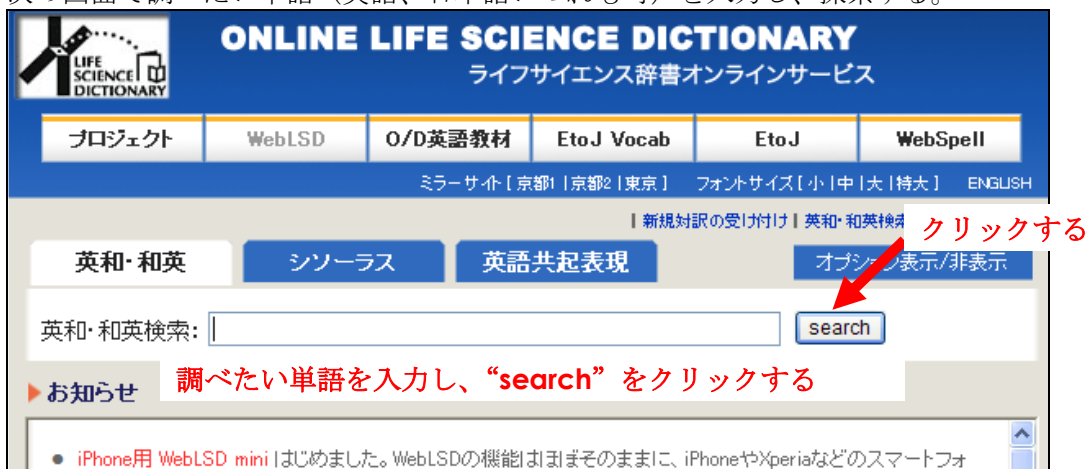
課題 1) 関心のある病気の選択を終えたら、その病気について、メルクマニュアル医学百科の記載事項で、特に重要と考えられる部分を Excel ファイルに記載をし、自分が設定した病気を十分に理解しておく。この際、自分の設定した疾患がウイルスや微生物による感染症でないことも再度確認をしておく。着目した疾患について、メルクマニュアル医学百科なかで、遺伝子が関係することを記載してある際は、その記載事項を Excel ファイルに記載せよ。以降の解析では、着目した疾患の英語名を知ること重要である。メルクマニュアル医学百科の記載事項中に英語名が出ていれば、それを Excel ファイルに記載せよ。併せて、ライフサイエンス辞書 (Google で検索せよ) のオンライン辞書サービスを用いて、疾患の英語名を調べて Excel ファイルに記載せよ。複数の英語名がある場合には、それらも、/ で区切って、Excel ファイルの同じセル内に記載せよ。メルクマニュアル医学百科中で記載されている英語名と一致するののかも調べよ。もし一致しない場合には、もう一度メルクマニュアル医学百科の記載事項を確認せよ。

ライフサイエンス辞書は、Google で“ライフサイエンス辞書”あるいは“LSD”というキーワードを入力し、検索ボタンをクリックする。次の画面が表示される。ライフサイエンス辞書プロジェクトをクリックする。

次の画面で、“WebLSD オンライン辞書” をクリックする。



次の画面で調べたい単語（英語、日本語いずれも可）を入力し、探索する。



ライフサイエンス辞書 <http://lsd.pharm.kyoto-u.ac.jp/ja/service/weblsd/index.html>

課題2) 次にその日本語の疾患名を Jabion 遺伝子百科 (Google で検索せよ) の検索の項目に入力し、その疾患に関する遺伝子を探索する。Jabion 遺伝子百科においては、その疾患ならびにその疾患に関する遺伝子の日本語での解説部分が存在するので、その部分を Excel ファイルに記載しておく。まずその疾患と関係をしている遺伝子が本来果たしている正常な働きを記載し、突然変異を起こして病気を引き起こすことについての記載があれば、その突然変異に関する事項も Excel ファイルに記載する。なお Jabion 遺伝子百科の検索結果を見ると、大半の病気について複数の遺伝子が関係していることが分かる。各々の遺伝子について、Excel ファイルの別の行に記載すること。

課題3) 次に、正常な機能をはたしている際の着目遺伝子の配列（言い換えれば、病気を引き起こす可能性のある突然変異を起こしていない配列）を Jabion 遺伝子百科に備わった機能で収集を行い、Excel ファイルに記載する。この際、以下の4種類の配列を収集する。

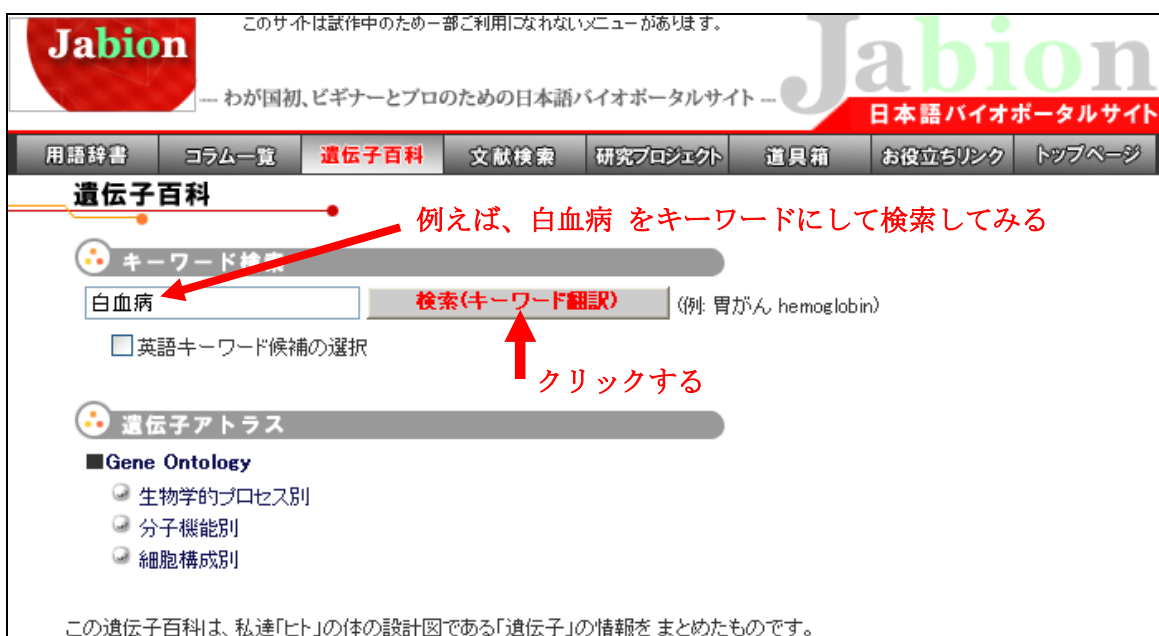
- (1) cDNA 配列
- (2) CDS (アミノ酸翻訳部分) の塩基配列
- (3) 遺伝子産物のタンパク質配列
- (4) ゲノム DNA 配列

なお、大半の病気について複数の遺伝子が関係している。各々の遺伝子について、上記の4配列を収集して、遺伝子ごとに Excel ファイルの別の行に記載する。

Jabion 遺伝子百科は、Google で“**遺伝子百科**”をキーワードに入力し検索する。次の様な結果画面が表示されるので、“**Jabion-日本語バイオポータルサイト- 遺伝子百科**”をクリックする。



次の画面が表示される。



Jabion 遺伝子百科 トップページ http://www.biportal.jp/Gene_search/search/search.cgi

自分の関心のある病気名を、(ここでは例えば“白血病”) キーワードに入力し検索ボタンをクリックすると、その病気に関連する遺伝子の一覧が表示される。(病気によっては Jabion 遺伝子百科では表示されない場合もある。)

このサイトは試作中のため一部ご利用になれないメニューがあります。

Jabion 日本語バイオポータルサイト

わが国初、ビギナーとプロのための日本語バイオポータルサイト

用語辞書 | コラム一覧 | **遺伝子百科** | 文献検索 | 研究プロジェクト | 道具箱 | お役立ちリンク | トップページ

遺伝子百科

キーワード検索

白血病 **検索(キーワード翻訳)** (例: 胃がん, hemoglobin)

英語キーワード候補の選択

検索結果:

生物種が **human** のものから、選択する。ここでは説明用として“**ABL1**”を選択した例で説明を進める。

遺伝子シンボル	遺伝子名	生物種	Link
ABCC5	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	human	NCBI
Abl1	v-abl Abelson murine leukemia oncogene 1	mouse	NCBI
ABL1	v-abl Abelson murine leukemia viral oncogene homolog 1	human	NCBI
Abl1_mapped	v-abl Abelson murine leukemia viral oncogene homolog 1 (mapped)	rat	NCBI
Abl2	v-abl Abelson murine leukemia viral oncogene 2 (arg, Abelson-related gene)	mouse	NCBI
ABL2	v-abl Abelson murine leukemia viral oncogene homolog 2 (arg, Abelson-related gene)	human	NCBI
Abl2_mapped	Abelson murine leukemia viral (v-abl) oncogene homolog 2 (mapped)	rat	NCBI
ANPEP	alanyl (membrane) aminopeptidase (aminopeptidase N, aminopeptidase M, microsomal aminopeptidase, CD13, p150)	human	NCBI
ANXA8	annexin A8	human	NCBI

“ABL1” をクリックすると次の画面が表示される。

Jabion **ABL1**

日本語バイオポータルサイト

Homologeneの相同遺伝子(HomoloGene:3783)

<input checked="" type="checkbox"/> H.sapiens ABL1 c-abl oncogene 1, receptor tyrosine kinase. NP_009297.2 1149 aa (NM_007313.2)	<input checked="" type="checkbox"/> P.troglodytes ABL1 c-abl oncogene 1, receptor tyrosine kinase. XP_001166213.1 1149 aa (XM_001166213.1)
<input checked="" type="checkbox"/> C.familiaris ABL1 c-abl oncogene 1, receptor tyrosine kinase. XP_548413.2 1131 aa (XM_548413.2)	<input checked="" type="checkbox"/> B.taurus ABL1 c-abl oncogene 1, receptor tyrosine kinase. XP_613548.2 1151 aa (XM_613548.4)
<input checked="" type="checkbox"/> M.musculus Abl1 c-abl oncogene 1, receptor tyrosine kinase. NP_033724.1 1123 aa (NM_009594.2)	<input checked="" type="checkbox"/> R.novaeigicus Abl1 c-abl oncogene 1, receptor tyrosine kinase. XP_001067860.1 1143 aa (XM_001067860.1)
<input checked="" type="checkbox"/> G.gallus ABL1	<input checked="" type="checkbox"/> D.rerio LOC100000720

画面を下にスクロールしていくと、次の画面が現れる。要約として、着目している遺伝子の機能や特徴等が日本語で記載されているので、この部分を読むこと。

H.sapiens *H.sapiens* (HAI) *P.troglodytes* *M.musculus* *R.norvegicus* *C.familiaris* *G.gallus* *D.laipes*

Basic Information

遺伝子シンボル: ABL1 (ABL, JTK7, bcr/abl, c-ABL, p150, v-abl)
 遺伝子名: c-abl oncogene 1

要約
 ABL1 (原癌遺伝子)は細胞質および核のチロシンリン酸化酵素を指令し、細胞「分化」、細胞分裂、細胞接着、およびストレス応答の過程に関係する。c-Abiタンパク質の活性はSH3領域によって抑制される。そして、SH3領域の「欠失」はABL1を「癌遺伝子」に変える。t(9;22)転座は慢性的な骨髄性白血病の多くの症例におけるBCR (MIM: 151410)とABL1遺伝子の直列融合をもたらす。広く発現されるABL1チロシン酸化酵素のDNA結合活性はCDC2の仲介するリン酸化によって調節され、ABL1の「細胞周期」機能を示している。ABL1遺伝子は6か7キロ塩基のmRNA転写物として発現され、選択的スプライシングにより一般的にエキソン2-11に第一エキソンが継がれる。

ゲノム配列上位置
 chromosome9
 base from 133589268 to 133763062 Chr. 9 133.59 M
 strand +
 extract sequence

エクソン・イントロン構造(クリックするとゲノムビューへジャンプします)
 NCBI
 Ensembl
 Nscan
 Ensembl

オリジナルDBIにおけるID・ACC(クリックするとオリジナルDBへジャンプします)
NCBI 25

no.	transcript ACC	peptide ACC
1	NM_005157.3	NP_005148.2
2	NM_007313.2	NP_009297.2

Ensembl ENSG00000097007

no.	transcript ACC	peptide ACC
1	ENST00000372348	ENSP00000361423
2	ENST00000393293	ENSP00000376971
3	ENST00000318560	ENSP00000323315
4	ENST00000426772	ENSP00000406308
5	ENST00000438426	ENSP00000407756
6	ENST00000444970	ENSP00000400412

オリジナルDBIにおけるID・ACC(クリックするとオリジナルDBへジャンプします)
NCBI 25
 Aligned RefSeq

no.	transcript ACC	peptide ACC
1	NM_007313.2	NP_009297.2

Ensembl ENSG00000097007

no.	transcript ACC	peptide ACC
1	ENST00000372348	ENSP00000361423
2	ENST00000393293	ENSP00000376971
3	ENST00000318560	ENSP00000323315
4	ENST00000426772	ENSP00000406308
5	ENST00000438426	ENSP00000407756
6	ENST00000444970	ENSP00000400412

ENSG00000097007 をクリックすると、全てのスプライシングバリエーションの配列情報を入手できる画面が表示される。

Ensembl west
 Home > Human [GRCh37] Login / Register | BLAST/BLAT | BioMart
 Location: 9:133,589,333-133,763,062 Gene: ABL1

Gene-based displays
 Gene summary
 Splice variants (6)
 Supporting evidence
 Sequence
 External references (3)
 Regulation
 Comparative Genomics
 Genomic alignments (52)
 Gene Tree (image)
 Gene Tree (text)
 Gene Tree (alignment)
 Orthologues (47)
 Paralogues (17)
 Protein families (4)
 Genetic Variation
 Variation Table
 Variation Image
 External Data
 Personal annotation
 ID History
 Gene history

Gene: ABL1 (ENSG00000097007)
 c-abl oncogene 1, receptor tyrosine kinase [Source:HGNC Symbol;Acc:76]
 Location Chromosome 9: 133,589,333-133,763,062 forward strand.
 Transcripts There are 6 transcripts in this gene: [hide transcripts](#)

Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
ABL1-001	ENST00000318560	5766	ENSP00000323315	1130	Protein coding	CCDS35166
ABL1-002	ENST00000372348	3824	ENSP00000361423	1149	Protein coding	CCDS35165
ABL1-003	ENST00000393293	532	ENSP00000376971	64	Protein coding	-
ABL1-201	ENST00000426772	144	ENSP00000406308	48	Protein coding	-
ABL1-202	ENST00000438426	3543	ENSP00000407756	265	Protein coding	-
ABL1-203	ENST00000444970	345	ENSP00000400412	115	Protein coding	-

Transcript and Gene level displays
 In Ensembl **Export data** をクリックする。
 Export data displays at two levels:
 • Transcript views which provide information specific to an individual transcript such as the cDNA and CDS sequence and domain annotation.
 • Gene views which provide displays for data associated at the gene level such as orthologues and paralogues, and splice variants.

次の画面が表示され、送ってもらうデータ (**Export data**) の内容や形式を指定できる。

Export Configuration - Feature List

Location to export: chromosome:GRCh37:9:133589333:133763062:1
 Gene to export: ENSG00000097007 (ABL1)
 Output: FASTA sequence
 Strand: Feature strand
 5' Flanking sequence (upstream): 0 * (Maximum of 1000000)
 3' Flanking sequence (downstream): 0 * (Maximum of 1000000)
 Next >

Fields marked with * are required

Options for FASTA sequence

Genomic: Unmasked
 cDNA
 Coding sequence
 Peptide sequence
 5' UTR
 3' UTR
 Exons
 Introns
 Fields marked with * are required

Options for CSV (Comma separated values)

Export Configuration - Feature List

Location to export: chromosome:GRCh37:9:133589333:133763062:1
 Gene to export: ENSG00000097007 (ABL1)
 Output: FASTA sequence
 Strand: Feature strand
 5' Flanking sequence (upstream): 0 * (Maximum of 1000000)
 3' Flanking sequence (downstream): 0 * (Maximum of 1000000)
 Next >

Fields marked with * are required

Options for FASTA sequence

Genomic: Unmasked
 cDNA
 Coding sequence
 Peptide sequence
 5' UTR
 3' UTR
 Exons
 Introns
 Fields marked with * are required

② クリックする

① 5' UTR, 3' UTR, Exons, Introns のチェックを外して [Next>] をクリックする

この画面で、Options for FASTA sequence の設定で、5' UTR, 3' UTR, Exons, Introns の チェックを外して [Next>] をクリックする。

Export Configuration - Output Format

Please choose the output format for your export

- HTML
- Text
- Compressed text (.gz)

< Back

Text をクリックする

次の画面が表示される。

```
>ENSG00000097007:ENST00000318560 cdna:KNOWN_protein_coding
GGCCTTCCCCCTGCGAGGATCGCCATTGGCCCCGTTTGGCTTTGCAAGCCGGCCCTGCGCT
TTGGGCCGGGCTCGGCCTCGGGAACGCCAGGGGCTC
GAGGGGGTTAAGGCGCAGGCGGGCGGGCGGGGGCGGGGG
( 中 略 )
```

↑ **cdna 配列**
cdna 配列をコピーして excel に貼り付ける。
コメント行 (" >" で始まる行) も含めてコピー

```
>ENSG00000097007:ENST00000318560 cds:KNOWN_protein_coding
ATGTTGGAGATCTGCCTGAAGCTGGTGGGCTGCAATCCAAGAAGGGGCTGTCCTCGTCC
TCCAGCTGTTATCTGGAAGAAGCCCTTCAGCGGCCA
GGTCTGAGTGAAGCCGCTCGTTGGAAGTCCAAGGAA
AATGACCCCAACCTTTTCGTTGCACTGTATGATTTT
AGCATAACTAAGGTCAGGCTCGGCTCTTAGGC
( 中 略 )
```

↑ **CDS 配列**
コピーして excel に貼り付ける
コメント行 (" >" で始まる行) も含めてコピー

```
>ENSG00000097007:ENST00000318560 peptide: ENSP00000323315 pep:KNOWN_protein_coding
MLEICLKLVGCKSKKGLSSSSCYLEEALQRPVASDIPDCTCFPAADPMNKEMTLAGDCE
NDPNLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEW
SLEKHSWYHGVPVSRNAAEYLLSSGINGSFLVRESESSPG
DCZTVMGCPDPMNTAFVWVHHCMTADCTTMTLVDADE
( 中 略 )
```

↑ **peptide 遺伝子産物のタンパク質配列**
コピーして excel に貼り付ける
コメント行 (" >" で始まる行) も含めてコピー

```
>ENSG00000097007:ENST00000372348 cdna:KNOWN_protein_coding
TGTGAGAATTGAAATGACTAGCATTATTGACCCTTTTCAGCATCCCCTGTGAATATTTCT
GTTTAGGTTTTTCTTCTTGAAAAGAAATTGTTATTCAGCCCGTTTAAAACAAATCAAGAA
ACTTTTGGGTAACATTGCAATTACATGAAATTGATAACCGCGAAAATAATTGGAAGTCCCT
GCTTGCAGCTCTCAACCTAABAAAAGCTGCTTCTTTCTTATGCAAGCATCTCTTTCTCTC
( 中 略 )
```

```
>ENSG00000097007:ENST00000372348 cds:KNOWN_protein_coding
ATGGGGCAGCAGCCTGGAAAAGTACTTGGGGACCAAAGAAGGCCAAGCTTGCCCTGCCCTG
CATTTTATCAAAGGAGCAGGGAAGAAGGAATCATCGAGGCATGGGGGTCCACACTGCAAT
GTTTTTGTGGAACATGAAGCCCTTCAGCGGCCAGTAGCATCTGACTTTGAGCCTCAGGGT
CTGACTCAAGCCGCTGCTGCAAGCTGCAAGCAAAAAGCTTCTGCTGCTGCAAGCAAGTCAAAAAT
( 中 略 )
```

```
>ENSG00000097007:ENST00000372348 peptide: ENSP00000361423 pep:KNOWN_protein_coding
MGQQPGKVLGDQRRPSLPALHFIRGAGKKESSRHGGPHCNVFEHEALQRPVASDFEPQG
LSEAARWNSKENLLAGPSENDPNLFVALYDFVASGDNTLSITKGEKLRVLGYNHNGEWCE
AQTKNGQGWVPSNYITPVNSLEKHSWYHGVPVSRNAAEYLLSSGINGSFLVRESESSPGQR
GCTDVMGCPDPMNTAFVWVHHCMTADCTTMTLVDADE
( 中 略 )
```

- (以下 **ABL1** の場合には、同じパターンで、
- ABL1-003 Transcript ID** [ENST00000393293](#) の **cdna, cds, peptide** の配列
 - ABL1-201 Transcript ID** [ENST00000426772](#) の **cdna, cds, peptide** の配列
 - ABL1-202 Transcript ID** [ENST00000438426](#) の **cdna, cds, peptide** の配列
 - ABL1-203 Transcript ID** [ENST00000444970](#) の **cdna, cds, peptide** の配列

が表示され、最後に **全ゲノム DNA 配列** が表示される。

```

>9 dna:chromosome chromosome:GRCh37:9:133589333:133763062:1
TGTGAGATTGAAATGACTAGCATTATTGACCCTTTTCAGCATCCCCTGTGAATATTTCT
GTTTAGGTCTTCTTCTTGAAAAGAAATTGTTATTCAGCCCGTTTAAAACAAATCAAGAA
ACTTTTGGG ゲノム DNA 配列 ABL1 の例 CTCCT
GCTTGCAAG “>” で始まる先頭行のみコピーして excel に貼り付ける CTGTG
ATTGACTTC ABL1 の例では 17 万文字程度の長い配列の為、excel の 1 マスに貼り
TTTCCCTCT 付けることが出来ない為、コメント行(“>” で始まる行)のみを excel GGGTA
CCTATTATT に貼り付ける。 GCCAA
GCTTGCCTGCCCTGCATTTTATCAAAGGAGCAGGGAAGAAGGAATCATCGAGGCATGGGG
GTCCACACTGCAATGTTTTTGTGGAACATGGTGAGTGCTTTTCAAATTTCTGCTCATGG
( 以下 略； ABL1 の例では 約 17 万文字の長い配列 )

```

このように、遺伝子 **ABL1** の場合は、

- ABL1-001 Transcript ID [ENST00000318560](#) の cdna, cds, peptide の各配列
 - ABL1-002 Transcript ID [ENST00000372348](#) の cdna, cds, peptide の各配列
 - ABL1-003 Transcript ID [ENST00000393293](#) の cdna, cds, peptide の各配列
 - ABL1-201 Transcript ID [ENST00000426772](#) の cdna, cds, peptide の各配列
 - ABL1-202 Transcript ID [ENST00000438426](#) の cdna, cds, peptide の各配列
 - ABL1-203 Transcript ID [ENST00000444970](#) の cdna, cds, peptide の各配列
- そして最後に、
- 全ゲノム DNA 配列
- 計 19 の配列情報が表示される。

ABL1 の例では、合計 **19** の各配列情報 (FASTA 形式) を、レポートの **excel** 様式の所定の欄に貼り付けることになる。

注) テキストの“ABL1”の例では 19 の配列情報であるが、調べる遺伝子によって、その数 (スプライシングバリエーションの数) は異なる。

注) FASTA 形式の配列では、行頭に “>” で始まるコメントの行があり、その次の行以降に配列が記述されている。各配列をコピーする時は、“>” で始まるコメントの行も含めてコピーすること。

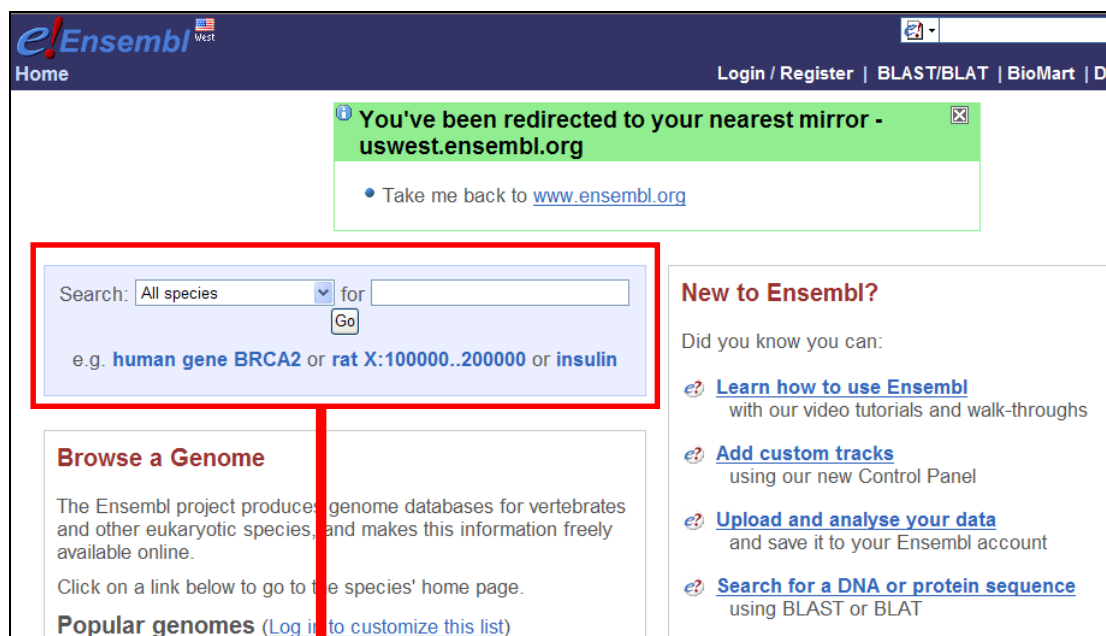
ここで、選択した疾患の日本語を Jabion 遺伝子百科で検索しても、遺伝子の情報が得られない場合も考えられる。その場合には、病気の英語名での検索も試みる。それでも着目の疾患に関する遺伝子情報が得られない場合には、Ensembl のような外国で作成しているゲノム配列情報に関するデータベースを直接に参照する必要がある。Ensembl の一般的な使用法は、統合 TV (Google で検索せよ) と呼ばれる動画サイトに、使用法が日本語で解説されている。その動画を参考にして Ensembl から上記の 4 種 (cDNA, CDS, peptide, 全ゲノム DNA) の配列を収集して、遺伝子ごとに Excel ファイルの別の行に記載する。但し、Ensembl を用いた配列の取得は、研究者が行う操作であり、Ensembl の機能が余りにも多いので、使用することは簡単ではない。13 と 14 頁にその操作を説明しているが、困難さを感じたら、Jabion で遺伝子が検索できるように設定した病気を変更しても良い。この際、以前に設定をして検索を行った結果は、この部分もレポートの評価の対象に含めるので、Excel ファイルからは除去しないこと。新たに設定した病名を、Excel ファイルの別の行に記載し、上記の操作を再開する。

※ Ensembl を使った配列データの取得

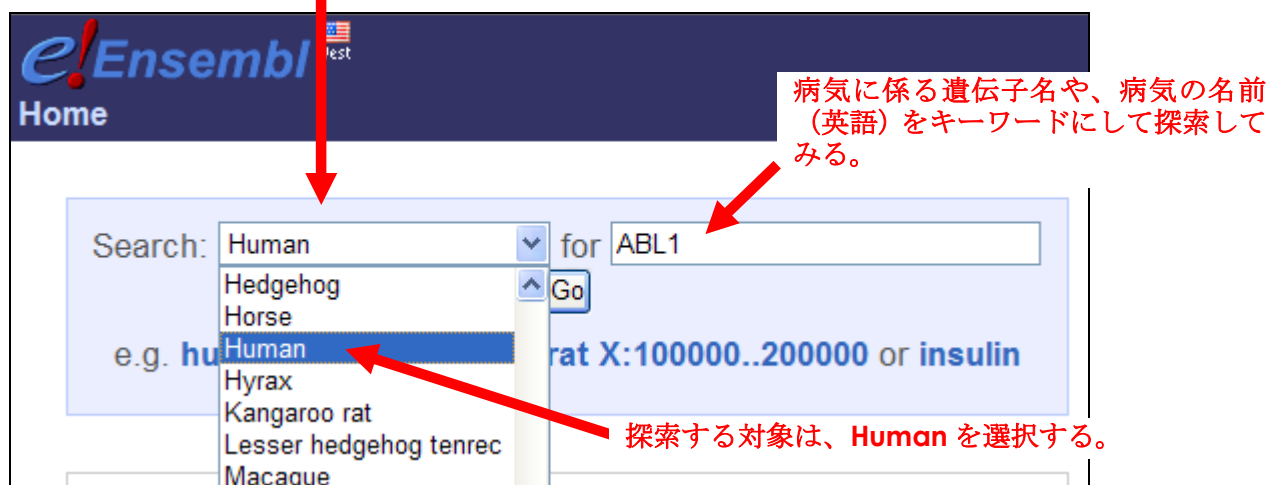
Google で “ensembl” をキーワードにして検索する。



Ensembl をクリックすると、次のトップ画面が表示される。



Ensembl トップページ (<http://uswest.ensembl.org/index.html>)



Ensembl 検索フォーム

Home > Human [GRCh37] Login / Register

Genome

Search Ensembl
 New Search
 Configure this page
 Manage your data
 Export data
 Bookmark this page

Results Summary

By Feature type

Total	5
Family	1
Gene	4
Homo sapiens (4)	

↑ クリックする

Home > Human [GRCh37] Login / Register | BLAST/BLAT | BioMart | Docs & Feedback

Genome

Search Ensembl
 New Search
 Configure this page
 Manage your data
 Export data
 Bookmark this page

Result in Detail

ABL1 corporate/tree:Top/Species/Homo sapiens c

Your query matched 4 entries in the search database

[Ensembl protein coding Gene: ENSG00000097007 \(HGNC Symbol: ABL1\)](#) [\[Region in detail\]](#)

Description: c-abl oncogene 1, receptor tyrosine kinase [Source:HGNC Symbol;Acc:76]

Source: e57; Feature type: Gene; Homo sapiens; Species: Homo sapiens; Gene;

[Havana protein coding Gene: OTTHUMG00000020813 \(ABL1\)](#) [\[Region in detail\]](#)

Description: c-abl oncogene 1, receptor tyrosine kinase

↑ クリックする

[Region in detail]をクリックすると、グラフィカルな画面が表示される。配列データ
 を取得する前に、こちらの画面も確認しておくとい。

Home > Human [GRCh37] Login / Register

Location: 9:133,589,333-133,763,062 **Gene: ABL1**

Gene-based displays

- Gene summary
 - Splice variants (6)
 - Supporting evidence
 - Sequence
 - External references (3)
 - Regulation
- Comparative Genomics
 - Genomic alignments (52)
 - Gene Tree (image)
 - Gene Tree (text)
 - Gene Tree (alignment)
 - Orthologues (47)
 - Paralogues (17)
 - Protein families (4)
- Genetic Variation
 - Variation Table
 - Variation Image
- External Data
 - Personal annotation
 - ID History
 - Gene history

Gene: ABL1 (ENSG00000097007)

c-abl oncogene 1, receptor tyrosine kinase [Source:HGNC Symbol;Acc:76]

Location [Chromosome 9: 133,589,333-133,763,062](#) forward strand.

Transcripts
 There are 6 transcripts in this gene:

Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
ABL1-001	ENST00000318560	5766	ENSP00000323315	1130	Protein coding	CCDS35166
ABL1-002	ENST00000372348	3824	ENSP00000361423	1149	Protein coding	CCDS35165
ABL1-003	ENST00000393293	532	ENSP00000376971	64	Protein coding	-
ABL1-201	ENST00000426772	144	ENSP00000406308	48	Protein coding	-
ABL1-202	ENST00000438426	3543	ENSP00000407756	265	Protein coding	-
ABL1-203	ENST00000444970	345	ENSP00000400412	115	Protein coding	-

Transcript and Gene level displays

In Ensembl a gene is made up of one or more transcripts. We provide displays at two levels:

- Transcript views which provide information specific to an individual transcript such as the c sequences and protein domain annotation.
- Gene views which provide displays for data associated at the gene level such as orthologous regulatory regions and splice variants.

This view is a gene level view. To access the transcript level displays select a Transcript ID in the navigate to the information you want using the menu at the left hand side of the page. To return to information click on the Gene tab in the menu bar at the top of the page.

↑ クリックする

この画面以降は、先程 P9 からと同じ。各配列をレポートの excel 様式に貼り付ける。

(参考) 統合TV (動画による Ensembl の操作説明)

Google で“統合TV 配列を取得する”をキーワードにして検索する。

The screenshot shows a Google search result for the query "統合TV 配列を取得する". The search results include a link to "統合TV (togotv)[塩基配列]" with a red arrow pointing to the word "クリックする" (click) in the snippet. Below the search results, there is a section titled "統合TV (togotv) - Ensembl tips ~配列を取得する~" with a red arrow pointing to the text "クリックするとFlash版 ..." (click to see the flash version ...).

The screenshot shows the "統合TV" website. At the top, there is a navigation bar with "統合TV" and a logo. Below the navigation bar, there is a search bar and a "クエリ送信" button. The main content area features a video player with a red arrow pointing to a video thumbnail titled "Ensembl Tip ~配列を取得する~". The thumbnail includes the text "090206 版" and a logo.

統合TV 画面 http://togotv.dbcls.jp/movie/090206en_f.html

II. 次に病気を引き起こす突然変異を知る操作に移る。

この場合には、OMIM (Online Mendelian Inheritance in Man ; ヒトの疾患データベース) を使用する。遺伝子が関係する病気に関する情報や、各病気を引き起こす遺伝子の突然変異を収録している、世界最大規模で最も信頼性の高いデータベースである。

課題4) 各自が選択した病気に関する遺伝子について、病気を引き起こすことに関係する遺伝子の突然変異のデータを収集する。タンパク質遺伝子について、病気と関係しているアミノ酸配列上の変化の位置とその変化の様子 (どのアミノ酸がどのアミノ酸へと変化したか等) を Excel に記載する。そして、余力のある学生はその記

載した突然変異体が、人種によって特徴があるかどうか、あるいはベンター博士やワトソン博士のゲノム塩基配列上に存在するのか等を調べてみる。

Google で“OMIM” をキーワードにして検索する。
次の画面が表示されるので、OMIM Home をクリックする。



OMIM トップページ <http://www.ncbi.nlm.nih.gov/omim/>

ここで、白血病に関連する“ABL1”の遺伝子を OMIM で調べてみる。for の右の入力ボックスに ABL1 と入力し Go のボタンをクリックすると、次の画面が表示される。

The screenshot shows the OMIM search interface. At the top, there is a search bar with 'OMIM' selected and 'for ABL1' entered. Below the search bar are navigation tabs: 'All Databases', 'PubMed', 'Nucleotide', 'Protein', and 'Genome'. A 'Limits' section shows 'All: 26' and 'OMIM dbSNP: 4'. The main results list shows three items, with the first item, '*189980. ABELSON MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1; ABL1', highlighted in purple. A red arrow points to the first item, and a red box highlights the text 'ABL1' in the title.

検索結果の各番号の次に表示されている記号について、*、+ は遺伝子に関するデータであることを示し、#、% は疾患に関するデータであることを示している。本自習では、遺伝子に関するデータに着目する。この場合、*印のついた ABL1 に関する遺伝子データが一番目にヒットしているので、これをクリックする。次の画面が表示される。

The screenshot shows the detailed view for MIM *189980. The title is 'ABELSON MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1; ABL1'. Under 'Alternative titles; symbols', it lists 'TRANSFORMATION GENE: ONCOGENE ABL', 'ABELSON STRAIN OF MURINE LEUKEMIA VIRUS; ABL', 'ABL1/BCR FUSION GENE, INCLUDED', and 'ABL1/NUP214 FUSION GENE, INCLUDED'. The 'Gene map locus' is '9q34.1'. On the right side, there is a 'Table of Contents' with a list of links: 'MIM *189980', 'Description', 'Cloning', 'Gene Structure', 'Mapping', 'Gene Function', 'Cytogenetics', 'Molecular Genetics', 'Animal Model', 'History', 'Allelic Variants', 'See Also', and 'References'. A red arrow points to the 'Allelic Variants' link.

この画面を下にスクロールしていくか、もしくは 右の Table of Contents の中の Allelic Variants をクリックすると次の様な Allelic Variants の情報が現れる。

次の画面で [See allelic variants as a separate report](#) をクリックする。

wrote the first chapter of the story of the molecular basis of the Philadelphia chromosome by reporting that the gene was translocated from chromosome 9 to chromosome 22.💡

Allelic Variants

クリックする

(Selected Examples) [Notes](#)

[Back](#)

[See allelic variants as a separate report](#)

.0001 LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB [ABL1, THR315ILE]

Gorre et al. (2001) reported a 944C-T transition in the region coding for the ATP-binding site of the ABL kinase domain of the BCR-ABL fusion product in 6 of 9 people with advanced-stage leukemia who were positive and imatinib-refractory. This mutation resulted in a thr315-to-ile (T315I) substitution. Thr315 forms an important hydrogen bond with the secondary amino group of imatinib (Schindler et al., 2000), suggesting that this mutation could be a major cause of resistance to imatinib.

.0002 LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB [ABL1, GLU255VAL]

In a patient with acute lymphoblastic leukemia (see 150555) positive for Philadelphia chromosome and

次の画面が表示される。1つの遺伝子の中に複数の変異の箇所があることがわかる。

Display Allelic Variants Show 20 Send to

All: 1 OMIM UniSTS: 0 OMIM dbSNP: 0

MIM *189980

ABELSON MURINE LEUKEMIA VIRAL ONCOGENE HOMOLOG 1; ABL1

Allelic Variants (Selected Examples) [Notes](#)

Number	Phenotype	Mutation	dbSNP
.0001	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, THR315ILE	-
.0002	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, GLU255VAL	-
.0003	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, GLU255LYS	-
.0004	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, TYR253HIS	-
.0005	CHRONIC MYELOID LEUKEMIA, RESISTANT TO IMATINIB	ABL1, PHE311LEU	-
.0006	CHRONIC MYELOID LEUKEMIA, RESISTANT TO IMATINIB	ABL1, MET351THR	-

“THR315ILE” の表記は、遺伝子産物のタンパク質配列の315番目のTHR（スレオニン）がILE（イソロイシン）に変異していると、発病の可能性が高いことを示している。この突然変異を起こしたアミノ酸配列情報を含むこの画面をExcelに記載しておく。

Number	Phenotype	Mutation	dbSNP
.0001	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, THR315ILE	-
.0002	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, GLU255VAL	-
.0003	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, GLU255LYS	-
.0004	LEUKEMIA, PHILADELPHIA CHROMOSOME-POSITIVE, RESISTANT TO IMATINIB	ABL1, TYR253HIS	-
.0005	CHRONIC MYELOID LEUKEMIA, RESISTANT TO IMATINIB	ABL1, PHE311LEU	-
.0006	CHRONIC MYELOID LEUKEMIA, RESISTANT TO IMATINIB	ABL1, MET351THR	-

↑ この枠内だけを、コピーし EXCEL に貼り付ける。

これまでの操作で、自分が着目した病気（この場合は白血病、LEUKEMIA）に関係するタンパク質遺伝子(ABL1)について、そのタンパク質のどの位置（複数の位置の場合も多い：ここでは、315 や 255、253 番目等のアミノ酸の位置）でどのようなアミノ酸変化が起こると、その病気になりやすいのかが調べられた。個人のゲノム配列が解読された際に、この Excel の表を参照すると、その人が罹り易い病気が推定できる。

今回の 遺伝子 ABL1 では、dbSNP（SNP に関するデータベース）へのリンク情報が無かったが、調べる遺伝子によっては、dbSNP（SNP に関するデータベース）へのリンク情報が記載されている場合がある。

例えば、白皮症に係る遺伝子（OCA2）について、このテキストの手順通りに調べると、次のような allelic variants as a separate report の画面が表示される。

SNP は“Single Nucleotide Polymorphism”の略で、日本語では一塩基多型と表現される。SNP について、このテキストでは説明を行わないが、各自 Google 等でキーワード検索して内容を調べ理解すること。

課題 5) 病気と関係する SNP データの取得を以下の操作に従って行う。

Number	Phenotype	Mutation	dbSNP
.0001	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, 2.7-KB DEL, EX7DEL	-
.0002	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, IVS17DS, G-T, +1	-
.0003	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, ALA481THR	rs74653330
.0004	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, VAL443ILE	rs28934272
.0005	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, PRO743LEU	-
.0006	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, 1-BP DEL	-
.0007	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, ALA334VAL	-
.0008	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, 122.5-KB DEL	-
.0009	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, TRP679CYS	-
.0010	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, ASN489ASP	-
.0011	SKIN/HAIR/EYE PIGMENTATION 1, BLUE/NONBLUE EYES	OCA2, ARG305TRP	rs1800401
.0012	SKIN/HAIR/EYE PIGMENTATION 1, BLUE/NONBLUE EYES	OCA2, ARG419GLN	rs1800407
.0013	SKIN/HAIR/EYE PIGMENTATION 1, BLUE/NONBLUE EYES	OCA2, IVS1, HAPLOTYPED 1	rs4778138,4778241,7495174
.0014	ALBINISM, OCULOCUTANEOUS, TYPE II	OCA2, MET394ILE	-

クリックする

全での SNP 情報を収集すること

dbSNP の各番号をクリックすると、NCBI ENTREZ の SNP データの画面が表示される。重要なデータであるので、dbSNP のデータがある場合は、必ず SNP の塩基配列情報も収集しておく。

②クリックする

(参考) : ここをクリックするとゲノムプロジェクトで得られた SNP 配列情報が表示される

① 配列をコピーして、excel に貼り付ける

rs74653330 をクリックすると次の画面が表示される。この中の Fasta sequence の配列をコピーして Excel に貼り付ける。

Single Nucleotide Polymorphism

Search Entrez SNP for [rs74653330] Go

Reference SNP(refSNP) Cluster Report: rs74653330 **clinically associated**

RefSNP	Allele
Organism: human (<i>Homo sapiens</i>)	Variation Class: SNP: single nucleotide polymorphism
Molecule Type: Genomic	RefSNP Alleles: C/T
Created/Updated in build: 131/132	Ancestral Allele: Not available
Map to Genome Build: 37.1	Clinical Association: VarView OMIM

SNP Details are organized in the following sections:

GeneView Map Submission Fasta Resource Diversity Validation

(中 略)

GeneView
 GeneView via analysis of contig annotation: [OCA2](#) *oculocutaneous albinism II*
 View more variation on this gene (click to hide).

Include clinically associated: in gene region cSNP has frequency double hit

Assembly	SNP to Chr	Chr	Chr position	Contig	Contig position	Allele
GRCh37	+	15	28228553	NT_026446.14	4663700	C

Function	mRNA				Protein		
	SNP to mRNA	Accession	Position	Allele change	Accession	Position	Residue change
missense	-	NM_000275.2	1551	GCC ⇒ ACC	NP_000266.2	481	A [Ala] ⇒ T [Thr]

(Open sequence viewer in a new window.)

(中 略)

Fasta sequence [\(Legend\)](#)

```
>gnl|dbSNP|rs74653330|allelePos=201|totalLen=401|taxid=9606|snpclass=1|alleles='C/T'|mol=Genomic|build=131
TGAGATGCCC AGTAGCACTT ACTGTGAAGA GGTGGCGTGA TGATCTTGAT TTCTAACAAA
GGCGTCCATG TGGGGTAGAG CTCTAACTAA GTGGAGGTGT GCGTTTACTG GAAGCAACCC
TAGCATGCTG GTACGTACCA TCTTCCTCAG CTCTTGTTG GAAAACAATAA TGACATTTGG
AGGGTCCCCG ATGGCAGTGG
Y
AGCTCCTCCA ATGTTTGTGA AGATCACTTC TGCAATCAGG ACTTGTCTTG GATCAAAGTT
GAGCACCTCA CACAACCTGT CACAAAATGGA GGAAAATGAA AGTAGTCCCA CTATACACAT
CGTGAAAAGGC CCACATGCAA CCCAGAGATT TTTAAGGCGC TTGCTCGTAT AAGGGAGGCG
CGCACACACA CACACCCTG
```

この配列を、コメント行 (“>” で始まる行) も含めてコピーして Excel の 1 マスの中に貼り付ける

NCBI Resource Links

Resource

Submitter-Referenced	dbSNP Blast Analysis	OMIM
GenBank		611409.0003
NC_000015		

Population Diversity

ss#	Sample Ascertainment				Genotypes	Alleles	
	Population	Individual Group	Chrom. Sample Cnt.	Source		C	T
ss127187185	pilot.1.CHB+JPT		88	AF		0.932	0.068

統計をとった北東アジア人集団 (CHB+JPT) では、93.2%が C、6.8%が T であったことを示している。 CHB (Chinese in Beijing 北京), JPT (Japanese in Tokyo)

数は少ないが、選んだ SNP によっては、Population の列にワトソン博士 (ENSEMBL Watson) やベンダー博士 (ENSEMBL Watson) の SNP 塩基配列情報が記載されている場合がある。調べた病気に係わる遺伝子の突然変異体がワトソン博士やベンダー博士の塩基配列上に存在するのを確認することも出来る。

課題6) 今回の実習で行った操作は、これからの医学において重要になる技術である。これまでの課題を行うことで操作の概略が理解できたと思う。比較的早く 課題5) までを終えた学生は、課題 3) で見つかる別の遺伝子について 課題 3) 以降のデータを収集し excel の別の行に追加すること。このことで今回の実習の全貌が理解できる。

(参考)

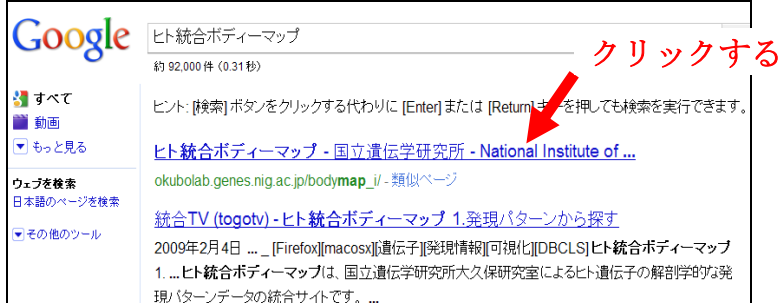
核酸塩基略号表	
略号	塩基 (略称の由来)
A	アデニン (Adenine)
T	チミン (Thymine)
G	グアニン (Guanine)
C	シトシン (Cytosine)
U	ウラシル (Uracil)
R	プリン (purine) A あるいは G
Y	ピリミジン (pyrimidine) T あるいは C
M	A あるいは C (amino)
K	G あるいは T (Keto)
S	G あるいは C (G と C の結合は強い (Strong))
W	A あるいは T (A と T の結合は弱い (Weak))
B	G あるいは T あるいは C (A の次は B)
H	A あるいは T あるいは C (G の次は H)
V	A あるいは G あるいは C (TU の次は V)
D	A あるいは G あるいは T (C の次は D)
N	AGTC のどれか (any)

※引用 : Wikipedia 核酸より

(参考)

ヒト統合ボディーマップは、国立遺伝学研究所大久保研究室が、文部科学省 統合データベースプロジェクトの成果の一つとして公開している。着目した遺伝子の発現パターンが確認出来る。統合TVに動画での操作説明があるので参考にするとよい。

Google で、“ヒト統合ボディーマップ” をキーワードにして探索をかける。



次のトップ画面が表示される。(http://okubolab.genes.nig.ac.jp/bodymap_i/)

画面右上のボックスに、例えば白血病に係わる遺伝子“ABL1”を入力して、検索ボタンをクリックする。

次の画面が表示される。

#	遺伝子ID	遺伝子名	最上位 遺伝子ファミリー	染色体上の位置	絶対発現量 (ピークカウント数/100万)			相対発現量 ■ 脳 ■ 血 ■ 結
					EST	SAGE NCBI 解釈	SAGE 大久保研 解釈	
1	Hs.431048	V-abl Abelson murine leukemia viral oncogene homolog 1	-	9q34.1(+)	167.1	207.3	133.5	未設計
2	Hs.11108	Cdk5 and Abl enzyme substrate 1	Cdk5 and c-Abl linker protein cables	18q11.2(+)	90.5	207.3	133.5	
3	Hs.301040	Cdk5 and Abl enzyme substrate 2	Cdk5 and c-Abl linker protein cables	20q13.33(-)	55.3	49.8	46.8	

■ 遺伝子情報

遺伝子 ID Hs431048
 ローカス ID 25
 遺伝子名 V-abl Abelson murine leukemia viral oncogene homolog 1
 同義遺伝子名 ABL, **ABL1**, Abelson murine leukemia viral oncogene homolog 1, JTK7, Proto-oncogene tyrosine-protein kinase **ABL1**, c-ABL, p150, v-abl
 メンバー構成 総数 : 597
 3' EST : 128
 5' EST : 291
 EST(Other) : 151
 HTC : 2
 mRNA : 22
 RefSeq : 2

iAFLPクラスタリング

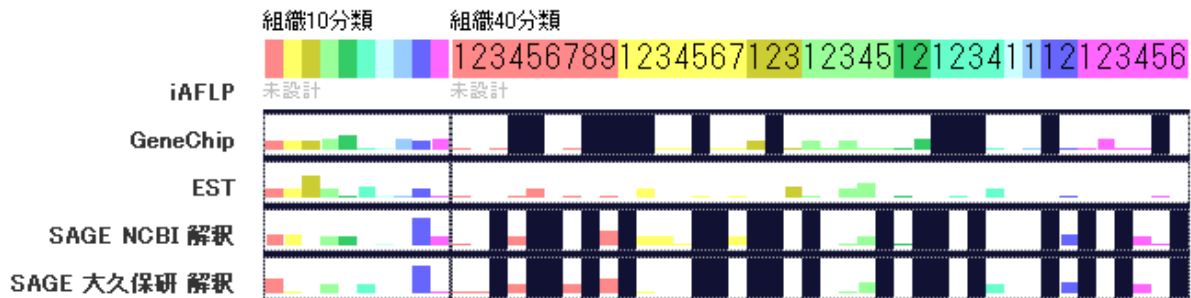
染色体位置 [9q34.1](#) 132,579,089~132,752,883(+)



■ 相対発現量

遺伝子の相対発現量です。

- グラフの高さは、相対発現量を表します。
- グラフの色は、臓器・組織を表します (内容は組織分類を参照)。
- xは、測定が失敗していることを表します (iAFLPのみ)。
- 紺の背景色は、臓器・組織が測定されていないことを表します。
- 10分類のグラフの高さは便宜上40分類のグラフの高さをカテゴリー(色)ごとに平均した高さです。



■ 組織分類

解剖学による臓器・組織の分類表です。

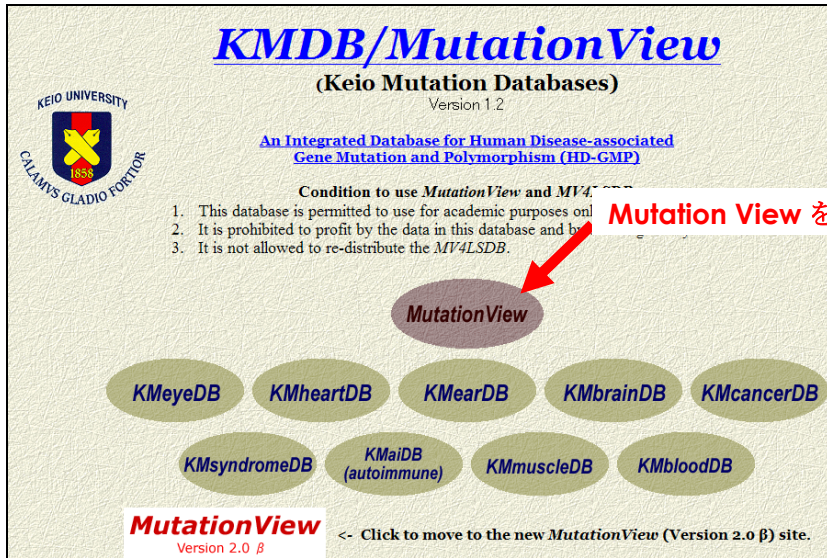
組織分類	1	2	3	4	5	6	7	8	9
10\40	1	2	3	4	5	6	7	8	9
脳	大脳	小脳	脳幹	脳梁	松果体	末梢神経	脊柱	網膜	目
血	動脈	静脈	リンパ節	末梢血	脾臓	胸腺	骨髄		
結	脂肪	骨	皮膚						
殖	胎盤	子宮	前立腺	卵巣	精巣				
筋	心臓	骨格筋							
消	食道	胃	腸	結腸					
肝	肝臓								
肺	肺								
腎	膀胱	腎臓							
泌	下垂体	甲状腺	副腎	膵臓	乳腺	唾液腺			

どの臓器・組織に相対的に発現量が多いかがわかる。

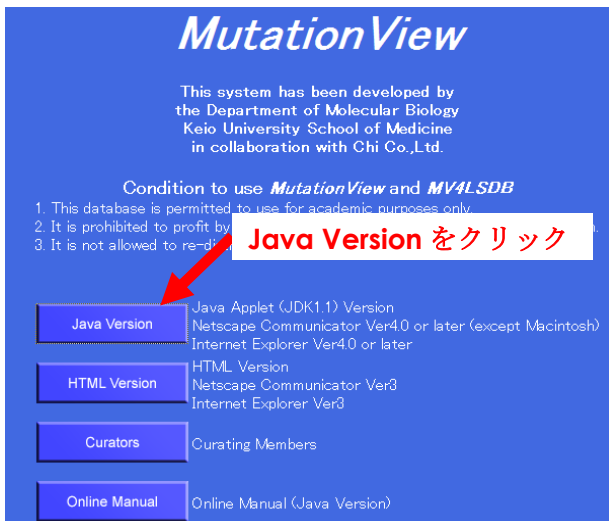
(参考)

余力のある人は、慶應大学（清水信義グループが作成）が公開している Mutation View(ミューテーションビュー)を参照することも、病気に関するデータベースへの理解を深める。なお、清水信義先生は長浜バイオ大の特別招聘教授です。

Google で “mutation view” をキーワードにして検索する。



<http://mutview.dmb.med.keio.ac.jp/mv/top>



MutationView

Java Version 1.3.57
Last Updated Apr.30, 2010

This system has been developed by
the Department of Molecular Biology
Keio University School of Medicine
in collaboration with Chi Co.,Ltd.

Chromosome Overview Chromosome Details **Anatomy**
OMIM Online Manual

Anatomy をクリック

Search by HUGO Gene Symbol OMIM Title OMIM Symbol

Searched 403 (895) now available
Gene(Disease): 413 (927) + being constructed

Gene: SELENOPROTEIN N, 1; SEPN1; SELN
NEPHROCYSTIN 4; NPHP4; KIAA0673; SNEPHRO
CHLORIDE CHANNEL, KIDNEY, B; CLCNKB; CLCK
PEROXISOME BIOGENESIS FACTOR 10; PEX10; PE
ENDOTHELIN-CONVERTING ENZYME 1; ECE1; EC
GAP JUNCTION PROTEIN, BETA-3; GJB3; GAP JUN
POTASSIUM CHANNEL, VOLTAGE-GATED, KQT-LIK
TUMOR-ASSOCIATED CALCIUM SIGNAL TRANSDU
Disease: Zellweger syndrome, 214100 (3); CEREBROHEPAT
Adrenoleukodystrophy, neonatal, 202370 (3); NEON
Hirschsprung disease, cardiac defects, and auton
[Hypertension, essential, susceptibility to], 145500 (3)
Hypophosphataemia, infantile, 241500 (3); HOPS; SP
Hypophosphataemia, infantile, 241500 (3)
Odontohypophosphatasia, 241500 (3); HYPOPHOS
Barter syndrome, type 3, 607364 (3); BARTTER SYN
Barter syndrome, type 4, digenic, 602522 (3); BSND
Nephronophthisis 4, 606966 (3); NEPHRONOPHTH

OMIM に比べると、登録されている疾患に係わる遺伝子数は少ないかもしれないが、優れた CG 技術を用いて身体の臓器別や病名別に病気に関係する遺伝子が表示されるので有用である。mutation view から疾患に係わる遺伝子を探索して、再び OMIM でその遺伝子の変異情報を取得してもよい。

(その他の参考サイト)

- Ensembl チュートリアル

<http://www-bird.jst.go.jp/minicourses/Ensembl1.html>

2010年6月1日初版

長浜バイオ大学 バイオサイエンス学部
コンピュータバイオサイエンス学科
生体分子情報学研究室 (池村研究室) 編集
URL : <http://www.nagahama-i-bio.ac.jp/>

※本テキストの内容等についてのお問い合わせは、
生体分子情報学研究室 上原 h_uehara@nagahama-i-bio.ac.jp 宛に願います。