

DNA 型鑑定と日本の科学捜査

(名古屋大学名誉教授)

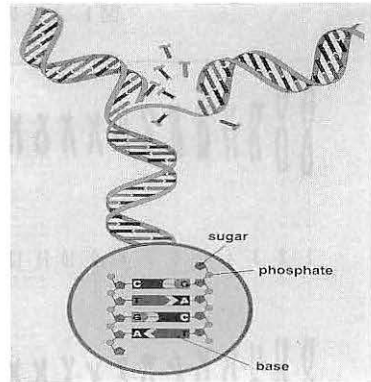
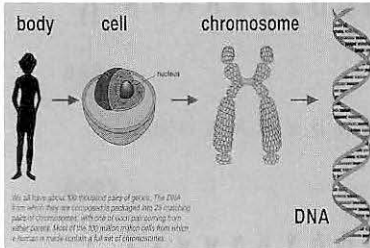
勝 又 義 直

1. DNA と遺伝情報

ただいまご紹介頂きました勝又です。今日は DNA 鑑定と日本の科学捜査ということで時間をいただきます。これは私が2005年に書いた本「DNA 鑑定 その能力と限界」ですが、この当時は既に DNA 鑑定が法廷でも使われていました。一方で、まだまだ結構議論もあったのですが、できるだけ DNA 鑑定の「信頼できますよ」という本です。副題は「その能力と限界」としているのですが、どちらかというと、能力があるので、きちんと使えば信頼できますよという本として書いております。それ以降、DNA 鑑定はかなり法廷で使われていくようになりましたが、多用されて



いくにつれていろいろな問題が出てきました。かなり難しい例で、無理筋の鑑定をすることもあるために、限界ではないかという例も出てきます。こちらの「最新 DNA 鑑定 その能力と限界」は、私が昨年改訂版として出した本ですけれども、どちらかというと限界についてご理解頂きたいということで、どこまでが可能でどこからが難しいのかというあたりを詳細に解説しま

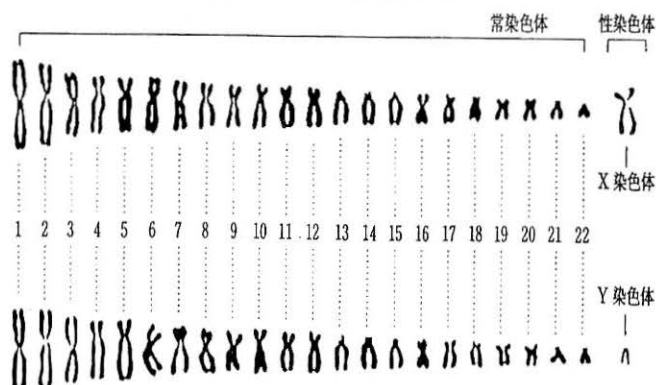


した。今日、パワーポイントでお見せしているのは、少しバージョンアップしたもので、ある程度の生データもご覧頂きながらお話したいと思います。

4点、お話ししていきますが、まずは基本情報です。DNAとは遺伝情報です。ヒトの遺伝情報は4種の塩基の配列で伝えられていますが、基本的な単位は30億塩基の連なりです。この30億の1基本単位をヒトゲノムといいまして、基本的には両親からそれぞれ一つゲノムが来ます。人類や多くの動物は、2倍体といいまして、対になった二つのゲノムによって身体の情報が出てきます。ですから我々の設計図は60億の塩基の連なりでできていることになります。我々はだいたい60兆の細胞でできていますが、基本は全部に核があってそこにあらゆる情報が入っています。

DNAの基本構造は二重螺旋でできていますが、4種類の塩基、シトシン (C)、グアニン (G)、アデニン (A)、チミン (T)、が情報を持って並んでいて、常に、二重鎖の一方のグアニンは他方のシトシンと、同じくアデニンはチミンとしか結合しません。そうすると、どういう利点があるかということ、複製がしやすい。つまり二重螺旋をはずして複製していくんですが、はずれた途端にその部分が二重になり、ダーッとはずれながら、

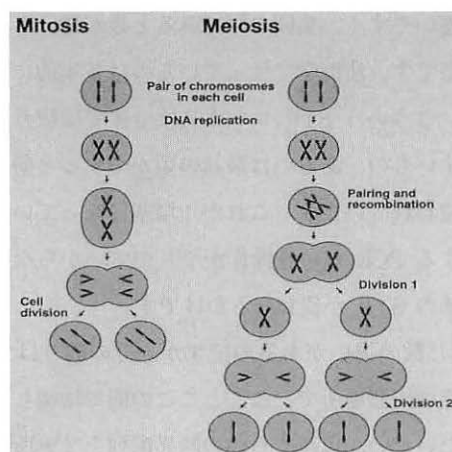
図1 ヒトの46本の染色体



染色体の端っことまで行きますと、二重螺旋が二つできているという、非常に効率の良い構造になっています。我々の遺伝情報はお父さん・お母さんそれぞれから一本ずつもらって46本の染色体で構成されています。このうち常染色体と言われるものが44本で、常にお父さんとお母さんと同じ形をしており大きいほうから番号が付いています。性染色体はX染色体とY染色体で、男性がXY、女性がXXの組み合わせになっています。

少し難しいですが、体細胞分裂と減数分裂の話をしてします。お父さんとお母さんの配偶子である精子と卵子が合体して一つの大元の受精卵ができ、その一つの細胞が分裂して我々の身体の60兆個の細胞ができるのですが、その分裂は全てこのミトース、いわゆる体細胞分裂で行われます。一つの細胞のDNAを二重にして、それが二つに分かれて二個の全く同じ細胞が作られる、いわゆるコピーです。これは間違いが少ないです。だいたい10万回分裂して1回くらい間違える程度の非常に安定した分裂です。ところが、配偶子を作る、つまり半分の情報を持つ精子と卵子を作るときには非常にややこしい分裂をします。これが、減数分裂で、それはいっぺん並ぶところまでは一緒なんです、そこで一度お父さんとお母さんの染色

(龍法 '15) 48-1, 592 (592)



体が少し情報交換をします。それがリコンビネーション、すなわち遺伝子組換えで、この操作があるために受精卵の配偶子は全てリセットされて全く新しい細胞に生まれ変わります。それが2つに分かれて最終的な4つの配偶子になるんですが、この一つ一つにお父さんとお母さんの遺伝子が少しずつ混じりあっています。この混じり合い方は様々で、そこで全く新しい能力が、もう一度生まれ変わるというような状態になります。これが減数分裂の重要なところ。この減数分裂はとてもややこしいので、突然変異が多いです。DNA 鑑定領域では平均0.2%、すなわち 500 に 1 回、間違えます。

2. DNA 鑑定の発展

DNA 鑑定は1985年から始まっていますが、現在主に使われているのは、short tandem repeat (STR) と言われる短い繰り返し配列の多型です。それはどういうものかという、DNA のいろんな部位で短い塩基の単位が繰り返している領域（ローカス）です。そういう領域では、減数分裂で

繰り返し数を間違いやすく、集団の中でみると数が違うものが共存しています。それが多型です。法医学で使っているのは基本的に4塩基単位の繰り返し配列です。こういうSTRローカスは全体では数万カ所あり、その中の一番使いやすいもの、あるいは最初の頃からずっと研究されているものが15ローカス使われています。これが4塩基になっている理由の一つは、試験管内で増幅するPCRという操作がやりやすいからなのです。そのためもっぱら4塩基のSTRが使われるわけです。

STRは繰り返し数の違いがありますが、その周りはすべて同じ塩基配列ですから、その同じ場所で、こことここ間で増幅しますというところを決めます。だいたい20塩基くらいの特異配列を2カ所決めれば、それらの間の領域を増幅します。この原理はなかなか難しいのですが、要するに温度を上げると二重螺旋がバラバラにはずれるようになります。そして、合成試薬が入ったまま温度をいっぺん下げると増幅します。ですから倍になるわけです。温度を更にまたもういっぺん上げるとまたバラバラになる。また更に温度を下げると、ワンサイクルでまた倍になる倍々ゲームなんです。合成には、温度が90℃以上に上げて壊れない酵素を使います。温泉などは90℃、100℃で湧いていますが、そこに住んでいる菌がそれだけの高温で平気なんですね。この酵素を使うアイデアを発明した人はノーベル化学賞を貰っています。効率良く温度を上げ下げするだけですごい勢いで倍々に増えていきますので、28回くらいそれをやると、だいたい10万から100万倍になる。たくさん同じものができますので分析が容易にできます。

どうやって分析するかと言いますと、長さが違ったDNA断片がいっぱいできるわけですから、それらを電気泳動で分離します。DNAは強い酸性です。核酸、つまりマイナスチャージがすごく強い。ですからプラスで引っ張ってやればどんどん流れていきます。その時にゲルという網目状物質の中でそれを動かします。そうすると、短いものは何にも遮られないの

でシューッと流れていきます。長いものはあちこちぶつかりながら動きますので、長さによって流れ方が違います。こういう原理を使って分析します。検出は通常、蛍光を用い、ゲルの中の位置がずれたところを検出すれば良いわけです。さらに、15 もの違う STR ローカスを増幅する試薬を一本の試験管の中に入れて一度に増幅し分析する技術が開発されました。マルチプレスキットというものですが、プレックスというのは増幅という意味で、たくさんのものを一度に増幅してしまう技術です。しかもそれを分析機械が自動吸引、自動型判定してくれます。

図 2

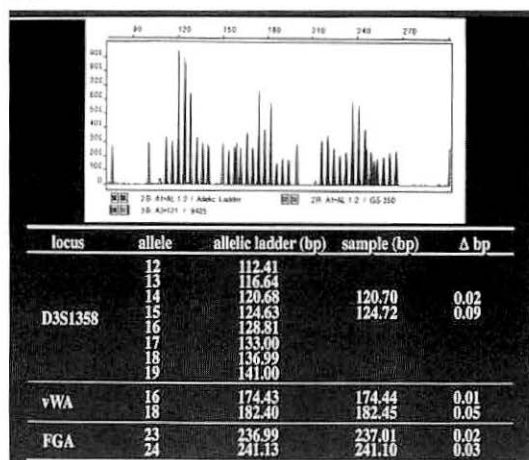


図 2 はアメリカのメーカーが開発したマルチプレックスの開発段階のものを調べた時のデータですが、グラフの横軸は 90 から 270 までの塩基対を示しているのですが、電気泳動すると短いのはビューと横軸左側まで流れます。長いのはあちこちぶつかって横軸右側のほうで止まってしまう。これを蛍光で測定します。グラフの縦軸は蛍光強度です。長さの具体的な評価は、長さの分かっている DNA を赤色蛍光で標識して入れてお

き、その位置を基準値にして計算します。そうすると、12回、例えばこのD3S1358 というローカスで12回繰り返している型は112塩基というふうに計算されます。計算ですから多少、小数点が出ますが、一応こういうふうに計算されます。そしてだいたい4塩基ずつ綺麗に型が並んでいます。なお、STR ローカスの型はこの繰り返し数を用いて表します。このローカスでは日本人の型はだいたい12から19出ますので、それを全部並べておくとこういうハシゴのようなものが出来る、横から見るとハシゴですから、アレリックラダーと言いますがこういうのが出来る。そして、誰かのものを入れると、例えば14と15のところにピッタリと一致して出てくる。この実際のものと基準としたラダーとの差が0.02塩基の長さしか変わらない。もともと型は4塩基ごとになっているのですから、だいたい0.2とか0.3塩基くらいしかズレがなく、非常に正確に型判定できることがわかります。

3. 現在の個人識別

個人識別には、ほぼ同じ手法が使われ続けています。感度を上げるなど改良は少しずつなされていますが原理は同じです。15ローカスのSTRを1本の試験管で一度に増幅して自動的に型判定します。去年頃からよく使われるようになった新しいキットは22ローカスに増えています。実はそれらのローカスはほとんどが別々の染色体に存在しており、一部は同じ染色体上にあるのですが場所がかけ離れていますので、基本的に独立していると考えて良いとなっています。ですから頻度の掛け合わせが可能です。一つひとつの検査ではそんなに明確な識別はできないのですが、それが全部一致しているかどうかというのが大きいと思います。

もう一つの特徴は、PCRは、もともとの存在量に比例して増幅するということです。同じ存在量だとほぼ同じ量が増幅されるという意味です。

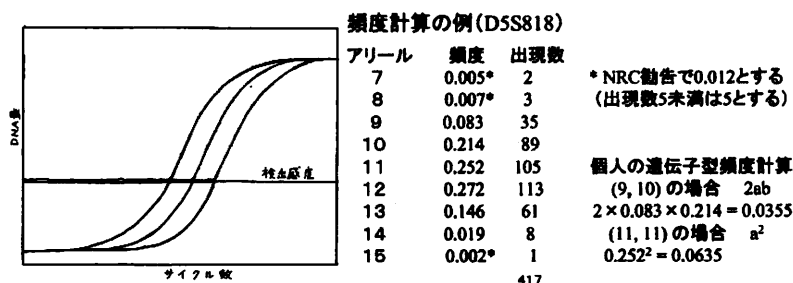
ですから微量な混入というのは非常に検出されにくく、ピークがものすごく低くてほとんど引かかってこないということです。通常の DNA 検査は28サイクルで、推奨されている DNA 量は0.5から2ng（ナノグラム）です。PCR 増幅の早い段階では微量すぎて何も検出できませんが、倍々ゲームで指数関数的に増えていき検出できるようになります。最後はヘタってしまっただけで止まりますが、この指数関数的に倍々ゲームで増えるところを使います。もともと倍量存在していたものはほぼ倍の高さのピークになるなど、元の DNA 量をほぼ反映して増幅されてくるのが基本的な仕組みです。15ローカスのそれぞれの系にはそれぞれの増幅効率があるので、それらをなるべく一度に測定しようということで、出やすいものはわざと出にくい条件に、出にくいものは出やすい条件にと、だいたい同じピークになるように調整しているのがキットです。キットを作るのにはかなり苦労します。私も6ローカスの特別のキットを作ったことがあるんですけども、うまく調整するのが大変でした。DNA の量にほぼ比例して増幅されるとどういうことが起こるかという、ホモ接合体ではお父さんとお母さんから同じものをもらっていますからピークの高さは倍になります。一方、ヘテロ接合体はホモ接合体の半分のピークになる。

そして、小さなピークを間違えて読まないように、ある高さ、閾値といいますが、以上のピークのみを読むことになっています。ですから現在の個人識別は、型判定をまず確実にしようという方針ですので、正確性はだいたい大丈夫です。それから再現性の確認もしようということになっています。同一 DNA について2回 PCR を行い、同じ結果になることを確認します。

それから、混入物の問題はなかなか厄介です。ヒト DNA はそこら中に溢れていますし、いったん増幅されてしまうと混入物と本来の資料 DNA の区別はできません。そのために微量な DNA はカウントしないようにします。例えば、倍々ゲームできちんと増えているところを使うと、混入

DNA はたいてい微量ですから検出しないだろうという考え方です。それでも大量の混入物がひょっとして入りかねないので、それを防止する操作マニュアルですね。手袋をしたり帽子を被ったり、あるいは滅菌の装置を使ったり、そういう操作を行います。

現在の出現確率の計算方法を一つのローカスを例にとって示します。一つのローカスの中でサンプル集団の中に現れる一つ一つの DNA の型をアリアルというのですが、このアリアルの出現頻度を調査すると、だいたいこのような山なりの頻度を示すことが多く、端っこは頻度が少ない。こういう場合、アメリカの NRC 勧告により出現数が2とか3など5以下は全部5と考えるというような配慮をします。小さな頻度の場合は、サンプル集団ごとでのバラツキが大きくなり、結果として被告人に不当に不利な計算結果になることがあります。それを防ぐために控えめな頻度を採用する配慮です。



ヒトはお父さんとお母さんからそれぞれのアリアルを受け継ぎますので、二つのアリアルの組み合わせになります。例にあげたお父さんとお母さんから9と10をもらった人の場合はその集団の中の出現頻度はだいたい3.5%と計算されます。そして11を両方からもらった場合は6.35%で、これだったら十数人に一人でしょうか。そのくらいの数字になります。一つひとつのローカスの出現確率だったらこの程度の判別しかできないので

表1 15種の STR を用いた個人識別の想定例①（新データベースによる修正）

ローカス	蛍光色素	B 血液	シャツ血痕	遺伝子型頻度	総合頻度
D8S1179	青	10,15	10,15	0.0356	0.0356
D21S11		28,30	28,30	0.0284	1.01E-03
D7S820		11,12	11,12	0.1543	1.56E-04
CFS1PO		12,12	12,12	0.1751	2.73E-05
D3S1358	緑	15,17	15,17	0.158	4.32E-06
TH01		9,9	9,9	0.1567	6.73E-07
D13S317		11,13	11,13	0.0226	1.53E-08
D16S539		9,12	9,12	0.1263	1.93E-09
D2S1338		21,23	21,23	0.0043	8.30E-12
D19S433	黄	12,13	12,13	0.0232	1.93E-13
vWA		17,19	17,19	0.0417	8.03E-15
TPOX		8,8	8,8	0.2065	1.66E-15
D18S51		13,17	13,17	0.0325	5.39E-17
Amerogenin	赤	X,Y	X,Y	—	—
D5S818		11,11	11,11	0.0854	4.60E-18
FGA		19,22	19,22	0.027	1.24E-19

(806京人に1人)

すけれども、ローカスはそれぞれが独立していますので全部が掛けられる。そうするとすごい総合出現確率になります。日本人などの集団内での総合出現確率は山形の分布になりますが、中央値としてはだいたい100～1000京人に一人になります。京は兆の1000倍多い単位です。日本人は1億人ちょっとしかいませんから、まあ、他にはなかなか出ないですね、というような感じになります。本に載せた想定事例は、個人が特定されてしまいますので、実際例を、ところどころ似たような確率の別の型に変えてあるのですけれども、ザッと計算してみますと、806京人に一人になります。

4. 微量 DNA 検査の問題点

現在、微量の分析がなんとかやれないかというのが研究課題になってお

りますので、少し説明します。我々は微量 DNA と言うのですが、「Low Copy Number」、LCN と略することが多いです。PCR に適した DNA 量は 0.5~2ng とされています。これだとだいたい80個から300個ぐらいの細胞相当ということになりますが、30個くらいまではなんとか安定したデータが出ます。それ以下になってきますと非常に不安定になり、この辺が限界です。どうして不安定になるかといいますと、初期段階の存在状態とか DNA が試薬とぶつかる確率がバラついてしまうからです。これを stochastic effect と言います。いわゆる確率的効果です。最初のバラツキがそのまま増幅されてしまうことになります。通常量の DNA でしたら、例えば二つのアリール、お父さんとお母さんの型があったとしますと、サンプルの中でもだいたい等量なのです。ですからある程度以上あれば、問題なくだいたい等量のピークが検出されます。それが、ものすごく微量になって DNA サンプルが数個の細胞分しかない場合には、お父さんのアリールばかり、あるいはお母さんのアリールばかり、なんてことがスタート時点で起こってしまう可能性があります。そうしますと、そのバラツキがそのまま増幅されてしまいます。本来同じ程度のピークが二つあるはずのヘテロ接合体でピークの高さが大きく違ったり、片方が消えてしまうなどいろんな事が起こってきます。

さらに困ったことに再現性が悪くなります。2回やって一致しないということです。これは確率の問題なので、なんともコントロールできない。例えば、同じサンプルを2回分析したら全然パターンが違いますとか、1回目は1本のピークしかないのに、もう1回やったら2本出てきましたとか、全然違うピークが出て来たり、ということが起きる。これはもう仕方がない部分ですが、やはり信頼性は低いです。なるべく投入量を多くしてそういうバラツキをなくそうとはしますが、難しい場合があります。こういう場合は PCR の感度をいくら上げても駄目です。最初がズレていますから。感度を上げればそれで確かにピークが高まるわけですけども、

元が不正確ですから、感度を上げるだけでは解決できない。確実性の保証が失われるということです。ですから、LCN の状態になれば出現確率の計算はしないというのが今のやり方です。

何とかならないかと、苦勞していろんな確率の計算を工夫したりしており、一カ所違っている程度なら、間違う確率を評価して確率に組み入れて出現確率を補正計算しようとの考えが出てきました。科学鑑定にもいろいろ問題がありますが、個人識別ならば本来は「すべてのローカスの型は必ず同じ」なのです。15 やれば 15、22 やれば 22 の型がすべて同じはずですよ。でも 15 やって 1 個しか違わないという、ものすごく惜しい訳ですよ。何かおかしいことが起こったのかも知れない、それを確率計算に入れようという考えがあって、ミスはこれくらいの確率で起こるのだから、1 つだけの違いであればそれを入れて計算してしまえというのが、今、ISFG、国際法遺伝学会の勧告でも出てきました。それでも今のところは、あちこちでバラバラ違っていたらもうお手上げ、と言うのが今の限界ではあります。観察されたピークが資料のものか混入したものかすらもわからなくなってくるのです。

それから注意が必要なのは、陳旧の斑痕資料です。古くなってきたものは核が壊れてしまって、細胞単位が存在が全く保証されません。バラバラになっていますから、そこを掬ってきても、例えば、お父さんとお母さんが一緒だから二つあるはずだ、同じ数あるはずだという保証が全然ありません。片方しかないなど、バラツキがひどくなっていく上に、やっぱり壊れてくるのですね。もしローカスの途中で切れていたらもうそこで PCR で増幅できませんので、全部揃っているものしか存在したことになりません。ですから元の DNA コピー数が多くても、残存 DNA が少ないと結果としては LCN となるという問題も生じます。

このように陳旧 DNA 資料は、十分量あってもうまく分析できない場合があります。分析に適した保存の問題というのは非常に難しく、新しい

資料でも、バクテリアが入って DNA を分解してしまうともう滅茶苦茶になります。一方、古いものでも、安定に保っていれば見える場合がないとは言えません。今のところは、せいぜい数十年くらいでしょうか。ものすごくいい状態で、凍結されている等ならば、何千年でもできないこともないかも知れません。それはケースによります。また、混合斑痕も少ない方の型は LCN になりやすいので、今盛んに研究されている分野です。

私からは以上です。どこまでが「限界」なのかについてお話しさせていただきました。

(石塚) 味噌樽に使っていたという資料などは、分析が難しいのでしょうか。

(勝又) やってみたいとわからないんですけれども、難しいと思いますね。結局、壊れやすいでしょうから。でも、ひょっとして保存はされていないとは言えませんし、あとはデータですね。データに再現性があれば大丈夫です。だから再現性がなかったり、出たり出なかったりとか、そういうような状況ですと、はっきりとしたことは言えない。一般的には、鑑定不能ですね。

(石塚) 鑑定不能と、そういうときには結論として出されますか。

(勝又) そういうことです。再現性があって、全部がそろっていたら鑑定可能です。

(石塚) どうでしょうか。いくつ駄目になったら鑑定不能というルールがあるのでしょうか。

(勝又) そうですね。それでも鑑定可能なんだということは言えるのかも知れないけれど。あんまり強くは言えないですね。