REVISTA DE MEDICINA

PUBLICAÇÃO DO DEPARTAMENTO SCIENTIFICO DO
CENTRO ACADEMICO "OSWALDO CRUZ" DA
FACULDADE DE MEDICINA DE SÃO PAULO - BRASIL

Director: SYLVIO BERTACCHI
Redactores: JAIRO CAVALHEIRO DIAS e ALOYSIO M. PIMENTA

VOL. XVIII | OUTUBRO — 1934 | NUM. 59

THEORIA ALGEBRICA DO CROSSING-OVER
CALCULO EXACTO DE SUA PROBABILIDADE

MARTINUS PAWEL
(DOUTORANDO)

No estudo da theoria do crossing-over, do entre-cruzamento chromosomico, da autoria de MORGAN, chama a attenção sobre si o facto de MORGAN medir as distancias entre os gens em porcentagens. Pois por varios motivos as porcentagens não se prestam á medição de distancias:

1.º — porque uma distancia é uma grandeza mathematicamente illimitada, emquanto uma porcentagem, quando empregada como neste caso, é limitada pelo numero 100 %; entretanto não é bem comprehensivel porque o comprimento dos chromosomas deva ter um limite tão preciso, ou por outro lado o que possa significar um comprimento chromosomico de p. ex. 118, correspondendo a 118 % de probabilidade de crossing-over. Porém, isto é de menor importancia visto que segundo o proprio MORGAN estes numeros maiores de 100, obtidos pela addição, não têm valor exacto, mas funccionam só como valores-limites superiores;

2.º — porque (e esta segunda objecção é muito mais séria) uma distancia sommada a outra distancia deve dar um resultado egual á somma arithmetica dos numeros que exprimem as distancias a sommar, e não, como acontece na somma de porcentagens, com desvios de 20 % e mais.

Estes motivos me levaram a repassar os trabalhos de MORGAN e a crear para elles um base mathematica, destinada a tornar possivel um calculo exacto dos indices de probabilidade de entre-cruzamentos chromosomicos e um calculo exacto tambem das distancias entre os gens — e não sómente uma avaliação approximada, como se obtinha apenas até hoje.

Entretanto necessario se torna salientar por causa da falta quasi absoluta de exemplos numericos neste terreno que para mais tarde, quando estes existirem em quantidade sufficiente, ha a possibilidade de não dar resultados satisfactorios o calculo com minhas formulas. Por este desvio não se deverá, porém, então culpar o methodo de calculo seguido, mas sómente uma de suas premissas, seja o principio fundamental abaixo assignalado, seja a propria theoria dos entre-cruzamentos de MORGAN, que então deverá ser corrigida ou substituida por outra. O principio fundamental e a hypothese dos entre-cruzamentos só se podem manter sob sua fórma actual se as minhas fórmulas derem resultados concordantes com os factos verificados.

Mas emquanto não possuirmos dados numericos e tabellas em numero sufficiente a respeito de addições de distancias entre gens todas criticas serão infructiferas.

Como principio fundamental da theoria mathematica a expôr admittimos (de accôrdo com MORGAN) a hypothese de que **a probabilidade de um entre-cruzamento entre 2 gens é proporcional á sua distancia.**

Precisamos, porém, divergir no ponto seguinte da maneira de pensar actualmente em voga, para não difficultar desde o inicio os calculos a effectuar: Não devemos dizer "em uma distancia **q** a probabilidade de p. ex. 2 entre-cruzamentos é p", mas sim que "em uma distancia **q** a probabilidade de p. ex. 2 **ou mais** entre-cruzamentos é p". Pois quando se calcula esta probabilidade segundo o principio fundamental acima relatado incluem-se no resultado todos os casos em que existem 2 entre-cruzamentos, existam ou não ainda outros entre-cruzamentos além destes 2. Isto se torna bem comprehensivel na deducção das fórmulas.

Por simplicidade exprimimos as probabilidades em cifras por unidade (/1), em lugar de exprimil-as em cifras por cento (%). As distancias entre os gens são expressas pela letra *z*, num systema de unidade que facilita o quanto possivel os calculos, excluindo todas constantes complicativas. As relações entre esta "unidade de entre-cruzamento" (U. E.) e a unidade Morgan serão postas a descoberto mais em baixo.

Deducção das fórmulas:

A probabilidade de 0 entre-cruzamentos ou mais (I. é de todos casos possiveis sem excepção) é naturalmente .... $s_0 = 1/1$

A probabilidade de 1 entre-cruzamento ou mais é, segundo o principio fundamental, quando se toma como constante de proporcionalidade a constante mais simples que ha, i. é o numero 1 .... $s_1 = z/1$

A prob. de 2 entrecr. ou mais é (como mostra fig. 1) $s_2 = z^2/1$

» » » 3 » » » » .... $s_3 = z^3/1$

« » » 4 » » » » .... $s_4 = z^4/1$

Generalizando, a probabilidade de n entre-cruzamentos ou mais é:

$$s_n = z^n \qquad (1)$$

Por meio desta fórmula podemos obter a probabilidade de cada entre-cruzamento isolado, e isto por simples subtracções:

A prob. de 0 entrecr. é .... $s_0 - s_1 = t_0 = 1 - z$

» » » 1 » » .... $s_1 - s_2 = t_1 = z - z^2$

» » » 2 » » .... $s_2 - s_3 = t_2 = z^2 - z^3$

etc.......

Generalizando novamente, vêmos que a probabilidade de um entre-cruzamento de ordem n é:

$$t_n = z^n - z^{n+1} = z^n . (1 - z) \qquad (2)$$

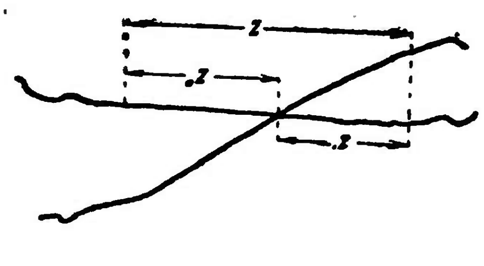

Fig. 1 — A probabilidade de 1 entrecruzamento no interior da distancia z e z, segundo o principio fundamental; a probabilidade de um segundo entrecruzamento é a fração expressa pela somma das listancias z' + z" = z da probabilidade z le um primeiro entrecruzamento, por conseguinte é z. (z' + z") = z2.

Como cada entrecruzamento impar entre 2 gens os separa, cada entrecruzamento par (tambem 0) os une, obtemos pela ddição das probabilidades t de todos os entrecruzamentos impaes a probabilidade v com que 2 gens se separarão, e pela addição e todas probabilidades de entrecruzamentos pares a probabiliade w com que permanecerão associados. Por conseguinte v é a probabilidade de um crossing visivel entre 2 gens, e **w** é a probabilidade dos crossings invisiveis (pares) ou inexistentes. Temos:

$$v = t_1 + t_3 + t_5 + t_7 + \dots \quad (3)$$

(2) $$v = z - z^2 + z^3 - z^4 + z^5 - z^6 + \dots \quad (4)$$

$$v = z.\,(1 - z + z^2 - z^3 + z^4 - z^5 + \dots) \quad (5)$$

e por outro lado:

$$w = t_0 + t_2 + t_4 + t_6 + \dots \quad (6)$$

(2) $$w = 1 - z + z^2 - z^3 + z^4 - z^5 + \dots \quad (7)$$

Mas conhecida fórmula mathematica nos dá:

$$z^0 - z^1 + z^2 - z^3 + z^4 - z^5 + \dots \pm z^n = \frac{1 - (-z)^{n+1}}{1 - (-z)} \quad (8)$$

e, como a razão *z* é menor do que *1*, pois é igual a uma probabilidade ($z = s_1$), temos: —

$$z^0 - z^1 + z^2 - z^3 + z^4 - z^5 \dots \pm z = \frac{1}{1 + z} \quad (9)$$

donde vem:

(5) $$v = \frac{z}{1 + z} \quad (10)$$

(7) $$w = \frac{1}{1 + z} \quad (11)$$

(10) e (11) $$v + w = 1 \quad (12)$$

Este ultimo resultado é natural, visto que a somma das probabilidades positivas e negativas de crossings visiveis deve ser a unidade. As fórmulas (10) e (11) dão a obtenção das probabilidades a partir da distancia entre os gens. Deduziremos agora as fórmulas inversas:

(10) $$v.\,(1 + z) = v + vz = z \quad (13)$$

$$v = z - vz = z.\,(1 - v) \quad (14)$$

$$z = \frac{v}{1 - v} \quad (15)$$

(11) $$w.(1+z) = w + wz = 1 \quad (16)$$
$$wz = 1 - w \quad (17)$$

$$z = \frac{1-w}{w} \quad (18)$$

A quantidade *z* é uma distancia que póde ser empregada em addições exactas; v e w são as probabilidades por unidade de separação e não-separação dos gens por um crossing.

Conhecendo 2 probabilidades de crossings visiveis com um gen commum póde-se pelas fórmulas precedentes calcular primeiro as duas distancias entre o gen commum e os dois outros (fórmula 15); depois, sommando algebricamente (somma arithmetica ou subtracção) as duas distancias, obtem-se a distancia entre os 2 outros gens e, a partir desta ultima, poder-se-á obter novamente uma probabilidade, a probabilidade de crossing entre estes 2 gens ultimos (fórmula 10). Isto se torna claro pelo exemplo abaixo.

Aliás podemos realizar os calculos com uma unica fórmula, caso só nos interessar o resultado final. Sejam z_1 e z_2 as distancias respectivas de 2 gens de um terceiro, v_1 e v_2 respectivamente as probabilidades de crossing entre elles e o terceiro gen commum, **Z** a distancia que medeia entre os dois primeiros gens e **V** a probabilidade de crossing entre elles, a determinar. Temos então:

15) $$z_1 = \frac{v_1}{1-v_1} \qquad z_2 = \frac{v_2}{1-v_2} \quad (19), (20)$$

(19) e (20) $$z_1 \pm z_2 = Z = \frac{v_1}{1-v_1} \pm \frac{v_2}{1-v_2} \quad (21)$$

Depois vem, substituindo o valor de **Z** em (21) na fórmula:

(10) $$V = \frac{Z}{1+Z} = Z:(1+Z) \quad (22)$$

$$V = \left(\frac{v_1}{1-v_1} \pm \frac{v_2}{1-v_2}\right) : \left(1 + \frac{v_1}{1-v_1} \pm \frac{v_2}{1-v_2}\right) \quad (23)$$

$$V = (v_1 \pm v_2 - v_1v_2 \mp v_1v_2) : (1 - v_2 \pm v_2 \mp v_1v_2) \quad (24)$$

Se ambas distancias se acharem no mesmo sentido, e portanto se devem ser sommadas arithmeticamente, obtem-se a probabilidade final:

$$V' = \frac{v_1 + v_2 - 2.v_1v_2}{1 - v_1v_2} \qquad (25)$$

Inversamente, se as duas distancias se acham collocadas em sentido contrario, e portanto devem ser subtrahidas uma da outra, obteremos:

$$V'' = \frac{v_1 - v_2}{1 - 2v_2 + v_1v_2} \qquad (26)$$

Exemplo e comprova:

Na **Drosophila melanogaster** temos em 33 % separação entre os caracteres "olhos brancos" (W) e "azas miniatura" (M) e em 22 % separação entre os caracteres "azas miniatura" (M) e "pés em furquilha" (F).

Baseado nas explanações de MORGAN só se póde inferir destes dados que a porcentagem provavel da separação entre W e F é menor que 55 % e maior que 11 % (isto é, menor que a somma e maior que a differença das duas probabilidades dadas.

Vejamos os resultados que obtemos com o emprego de nossas fórmulas:

Dados: — $v_1 = 33\% = 0{,}33/1$ $\quad v_2 = 0{,}22\% = 0{,}22/1$

Por conseguinte devemos ter, segundo (15):

$$z_1 = \frac{0{,}33}{1 - 0{,}33} = 0{,}4925$$

$$z_2 = \frac{0.22}{1 - 0{,}22} = 0{,}2820$$

Estas distancias devem estar situadas ou no mesmo sentido ou então em sentido contrario; no primeiro caso deverão ser sommadas, no segundo caso deve ser subtrahida uma da outra:

$$Z' = z_1 + z_2 = 0{,}4925 + 0{,}2820 = 0{,}7745$$

$$Z'' = z_1 - z_2 = 0{,}4925 - 0{,}2820 = 0{,}2105$$

Estas distancias podem agora servir, segundo a fórmula (10), para o calculo da probabilidade de entrecruzamentos visiveis:

$$V' = \frac{Z'}{1 + Z'} = \frac{0{,}7745}{1{,}7745} = 0{,}4365/1 = 43{,}65\ \%$$

$$V'' = \frac{Z''}{1 + Z''} = \frac{0{,}2105}{1{,}2105} = 0{,}1740/1 = 17{,}40\ \%$$

Realmente o numero obtido por contagens directas é ca. 44 %, coincidindo bem com V'; a differença decimal póde ser facilmente explicada pelo arredondamento da porcentagem e pelas fluctuações naturaes.

Fazendo o calculo approximado por meio das fórmulas directas (25) e (26) obtem-se:

$$V' = \frac{0{,}33 + 0.22 - 2 \times 0.33 \times 0.22}{1 - 0{,}33 \times 0{,}22} = \frac{3 \times 9 + 2 \times 9 - 2 \times 3 \times 2}{81 - 3 \times 2} =$$

$$= \frac{33}{75} = \frac{44}{100} = 44\ \%$$

$$V'' = \frac{0{,}33 - 0{,}22}{1 - 2 \times 0{,}22 + 0{,}33 \times 0.22} = \frac{1 \times 9}{81 - 2 \times 2 \times 9 + 3 \times 2} =$$

$$= \frac{9}{51} = \frac{18}{102} = 17{,}64\ \%$$

Se fizessemos o calculo exacto, naturalmente os resultados seriam exactamente os mesmos que os obtidos anteriormente.

Como os outros dados numericos publicados que pudemos consultar ainda são muito pouco exactos por se basear em estatisticas muito pequenas, excedendo mesmo os limites estabelecidos por MORGAN (v. TABULAE BIOLOGICAE, ed. Junk), não me é possivel apresentar outros exemplos exactos.

Erros menores, principalmente os notados entre gens proximos, podem ser explicados pelo facto de que os chromosomas não são linhas geometricas, mas possuem certa espessura, pela qual as distancias podem ser algo augmentadas.

Como se define minha nova unidade de distancia entre os gens, unidade de entrecruzamento? (U. E.) É muito simples conseguir tal definição. Façamos z = 1 na fórmula (10). obteremos:

$$v = \frac{z}{1 + z} = \frac{1}{1 + 1} = 1/2 = 0{,}5/1 = 50\ \%$$

Por conseguinte, a minha U. E. é a distancia necessaria para dar 50 % de entrecruzamentos visiveis, e eo ipso tambem 100 — 50 = 50 % de ajoujamentos.

Como um Morgan é a distancia na qual a probabilidade de separação é 1 % ou 0,01/1, temos:

$$z = \frac{0,01}{1 - 0,01} = \frac{1}{99}$$

Quer dizer isto que minha U. E. contem 99 morgans. Como porém o morgan não é uma constante indicadora da distancia, mas uma variavel (quando z =1, v = 50 %, portanto aqui 1 U. E. só equivaleria a 50 morgans), este numero não póde ser utilizado para a conversão. Devem ser usadas para tal as fórmulas supra-indicadas (11) e (16) ou a curva da figura 2, com cujo auxilio se podem converter morgans ou probabilidades (em % ou /1) em U. E. e vice-versa, e por meio da qual tambem é possivel resolver graphicamente os calculos a fazer em casos como o exemplificado.

NOTA — Este trabalho foi publicado já quasi nos mesmos termos na Zeitschrift fuer induktive Abstamungs und Vererbungslehre, vol. LXV, pag. 141. Approveitei sómente esta publicação em portuguez para corrigir alguns erros typographicos que, embora não affectavam o resultado final dos calculos, difficultavam sua comprehensão.

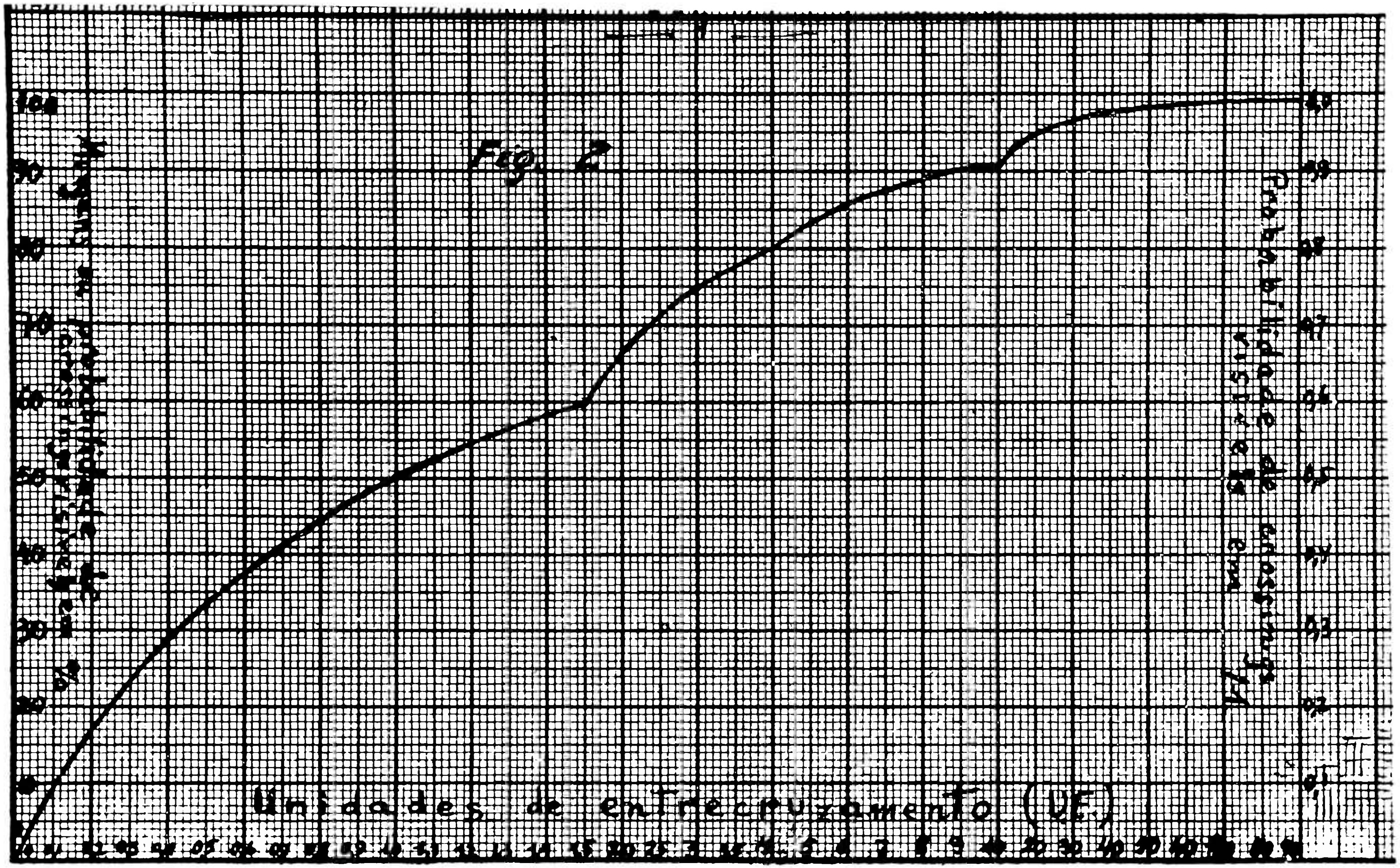
Fig. 2
Unidades de entrecruzamento (UE)
Probabilidade de crossings visíveis em %
Margens da probabilidade de crossings visíveis em %