



Programa de Pós-Graduação em
Biologia Marinha e Ambientes Costeiros
UFF

UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE BIOLOGIA
CURSO DE PÓS GRADUAÇÃO EM BIOLOGIA MARINHA E
AMBIENTES COSTEIROS

**ESTUDOS EM GENÉTICA MARINHA E EVOLUÇÃO: VARIAÇÃO
GÊNICA, MORFOMETRIA MULTIVARIADA, ESTRUTURA DE
POPULAÇÕES E PADRÕES DE BIODIVERSIDADE**

VOLUME I

MICHELLE REZENDE DUARTE

NITERÓI

OUTUBRO 2016



Programa de Pós-Graduação em
Biologia Marinha e Ambientes Costeiros
UFF

UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE BIOLOGIA
CURSO DE PÓS GRADUAÇÃO EM BIOLOGIA MARINHA E
AMBIENTES COSTEIROS

**ESTUDOS EM GENÉTICA MARINHA E EVOLUÇÃO: VARIAÇÃO
GÊNICA, MORFOMETRIA MULTIVARIADA, ESTRUTURA DE
POPULAÇÕES E PADRÕES DE BIODIVERSIDADE**

VOLUME I

MICHELLE REZENDE DUARTE

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biologia Marinha e Ambientes Costeiros, Departamento de Biologia Marinha, Instituto de Biologia, Universidade Federal Fluminense, como requisito parcial para a obtenção do Título de Doutor em Biologia Marinha e Ambientes Costeiros.

ORIENTADOR:
DR. EDSON PEREIRA DA SILVA

NITERÓI
OUTUBRO 2016

D 812 Duarte, Michelle Rezende

Estudos em genética marinha e evolução: variação gênica, morfometria multivariada, estrutura de populações e padrões de biodiversidade/Michelle Rezende Duarte. - Niterói: [s. n.], 2016.

3v. (xxv; 687f.).

Tese – (Doutorado em Biologia marinha e Ambientes Costeiros) – Universidade Federal Fluminense, 2016.

Inclui v.1 e v. 2: Conteúdo textual.

Inclui v.3: Apêndices.

1. Genética populacional. 2. Evolução. 3. Bioinvasão. 4. Biodiversidade. 5. Morfometria. 6. Actinopterygii. 7. Bivalve. 8. Inventário. I. Título.

CDD.: 591.15

**ESTUDOS EM GENÉTICA MARINHA E EVOLUÇÃO: VARIAÇÃO GÊNICA,
MORFOMETRIA MULTIVARIADA, ESTRUTURA DE POPULAÇÕES E
PADRÕES DE BIODIVERSIDADE**

VOLUME I

MICHELLE REZENDE DUARTE

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biologia Marinha e Ambientes Costeiros, Departamento de Biologia Marinha, Instituto de Biologia, Universidade Federal Fluminense, como requisito parcial para a obtenção do Título de Doutor em Biologia Marinha e Ambientes Costeiros.

BANCA EXAMINADORA:

DR. EDSON PEREIRA DA SILVA (PRESIDENTE)
UNIVERSIDADE FEDERAL FLUMINENSE

DR. AGUINALDO NEPOMUCENO MARQUES JUNIOR
UNIVERSIDADE FEDERAL FLUMINENSE

DR. CASSIANO MONTEIRO NETO
UNIVERSIDADE FEDERAL FLUMINENSE

DR. CRISTIANO VALENTIM DA SILVA LAZOSKI
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

DRA. GISELLE LOBO HAJDU
UNIVERSIDADE ESTADUAL DO RIO DE JANEIRO

DR. PAULO CESAR DE PAIVA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

DR. RAFAEL DE ALMEIDA TUBINO
UNIVERSIDADE FEDERAL FLUMINENSE

*Quando o mar
Quando o mar tem mais segredo
Não é quando ele se agita
Nem é quando é tempestade
Nem é quando é ventania
Quando o mar tem mais segredo
É quando é calmaria*

Amor Amor (1977)
Sueli Costa & Cacaso

À minha mãe, que permitiu que eu buscasse os meus sonhos e ao meu companheiro de vida acadêmica, que acreditou neles, caminhou ao meu lado e ajudou a realizá-los.

AGRADECIMENTOS

Agradeço à minha base, minha amada família, por todo carinho e dedicação. Minha Mãe, a melhor mãe do mundo, que sempre me apoiou e esteve ao meu lado em todos os momentos da minha vida, lutando comigo para transformar meus sonhos em realidade. Meu Pai_rasto, Paulo Roberto, por toda força, sempre e ao meu irmãozinho (?) Sávio, por existir. Minha tia Elenice e meus tios Elbson e Manuel, por serem tão presentes na minha vida e me darem tanto amor. À Mari, minha amiga mais que amiga, minha irmã de coração, companheira de todas as horas, que tem estado ao meu lado há muitos anos.

A todos do Coletivo do Laboratório de Genética Marinha e Evolução, muito obrigada pelo companheirismo, pelas conversas, pelos momentos que tivemos que falar sério, pelos momentos de descontração, pelas nossas reuniões. Vocês foram muito importantes.

À Ana Cristina por sempre me receber com muita atenção, carinho e paciência. Um agradecimento também a todos os professores do PBMAC, por todo processo de aprendizagem que passei e a CAPES, pela bolsa concedida.

A todos aqueles amigos especiais que fizeram parte deste trabalho (direta ou indiretamente).

E, finalmente, ao Edson, que foi incrível, um grande companheiro neste caminho. Obrigada por tudo, orientador!

ÍNDICE VOLUME I

AGRADECIMENTOS.....	vi
ÍNDICE VOLUME II.....	ix
ÍNDICE VOLUME III.....	xi
LISTA DE FIGURAS VOLUME I.....	xii
LISTA DE FIGURAS VOLUME II.....	xvi
LISTA DE QUADROS VOLUME I.....	xviii
LISTA DE TABELAS VOLUME I.....	xx
LISTA DE TABELAS VOLUME II.....	xxi
RESUMO.....	xxiii
ABSTRACT.....	xxiv
INTRODUÇÃO GERAL.....	25
OBJETIVOS.....	28
1. GENÉTICA MARINHA.....	30
1.1. INTRODUÇÃO.....	31
<i>1.1.1. Genética de Populações.....</i>	<i>32</i>
<i>1.1.2. Biologia Marinha.....</i>	<i>34</i>
<i>1.1.3. Genética Marinha.....</i>	<i>36</i>
1.2. HISTÓRICO.....	38
<i>1.2.1. Antes do Começo.....</i>	<i>39</i>
<i>1.2.1.1. A teoria evolutiva darwiniana.....</i>	<i>40</i>
<i>1.2.1.2. A teoria sintética da evolução.....</i>	<i>45</i>
<i>1.2.1.3. Controvérsias evolutivas.....</i>	<i>50</i>
<i>1.2.2. O Começo.....</i>	<i>54</i>
<i>1.2.2.1. No mundo.....</i>	<i>55</i>
<i>1.2.2.2. No Brasil.....</i>	<i>59</i>
<i>1.2.2.3. Nos outros.....</i>	<i>63</i>
1.3. EMPÍRICO.....	65
<i>1.3.1. Variação Fenotípica.....</i>	<i>67</i>
<i>1.3.2. Morfometria.....</i>	<i>73</i>
<i>1.3.3. Variação Gênica.....</i>	<i>77</i>
<i>1.3.4. Citogenética.....</i>	<i>81</i>
<i>1.3.5. Eletroforese de Aloenzimas.....</i>	<i>84</i>
<i>1.3.5.1. A técnica.....</i>	<i>85</i>
<i>1.3.5.2. O método.....</i>	<i>87</i>
<i>1.3.6. Técnicas de DNA.....</i>	<i>90</i>
<i>1.3.6.1. DNA alvo.....</i>	<i>92</i>
<i>1.3.6.2. Tipos de sequências.....</i>	<i>95</i>
<i>1.3.6.3. RFLP.....</i>	<i>97</i>
<i>1.3.6.4. PCR.....</i>	<i>99</i>
<i>1.3.6.5. Sequenciamento.....</i>	<i>104</i>
<i>1.3.7. Ómicas.....</i>	<i>106</i>
<i>1.3.8. E Agora José?.....</i>	<i>109</i>
1.4. CONCEITUAL.....	116
1.4.1. Variação Intrapopulacional.....	118
<i>1.4.1.1. Descrição dos dados empíricos.....</i>	<i>118</i>
<i>1.4.1.2. Testes de hipóteses.....</i>	<i>123</i>
<i>1.4.1.3. Em nível de sequências.....</i>	<i>130</i>
1.4.2. Variação Interpopulacional.....	134
<i>1.4.2.1. Estrutura populacional.....</i>	<i>134</i>

<i>1.4.2.2. Modelos.....</i>	141
<i>1.4.2.3. Relações genéticas.....</i>	145
1.4.3. Relógio Molecular.....	150
<i>1.4.3.1. Taxa de substituição.....</i>	151
<i>1.4.3.2. Taxa das sequências.....</i>	155
<i>1.4.3.3. Taxa dos taxa.....</i>	157
<i>1.4.3.4. Pontos de calibração.....</i>	159
1.4.4. Coalescência.....	160
1.4.5. Filogeografia.....	162
1.5. COMPUTACIONAL.....	166
<i>1.5.1. Revolução Computacional.....</i>	167
<i>1.5.2. Softwares.....</i>	171
<i>1.5.3. Em Frente para o Passado.....</i>	173
1.6. CIENTÍFICO.....	175
<i>1.6.1. Variação Gênica.....</i>	178
<i>1.6.2. Estruturação.....</i>	189
<i>1.6.3. Sistemática e Taxonomia.....</i>	202
<i>1.6.4. Genética Ecológica.....</i>	210
<i>1.6.5. Genética e Poluição.....</i>	215
<i>1.6.6. Genética Pesqueira.....</i>	219
<i>1.6.7. Aquicultura.....</i>	224
<i>1.6.8. Bioinvasão.....</i>	228
<i>1.6.9. Conservação.....</i>	234
<i>1.6.10. Biogeografia.....</i>	241
<i>1.6.11. Microbiologia.....</i>	246
1.7. CONCLUSÃO.....	250

ÍNDICE VOLUME II

LISTA DE FIGURAS.....	cclix
LISTA DE TABELAS.....	cclxi
2. RECURSOS PESQUEIROS.....	263
2.1. INTRODUÇÃO.....	264
2.2. DOURADO.....	267
2.2.1. Material e Métodos.....	269
2.2.1.1. Área de estudo.....	269
2.2.1.2. Eletroforese de aloenzimas.....	270
2.2.1.3. Análise estatística.....	273
2.2.2. Resultados.....	277
2.2.3. Discussão.....	285
2.3. XERELETE.....	290
2.3.1. Material e Métodos.....	293
2.3.1.1. Área de estudo.....	293
2.3.1.2. Morfometria geométrica.....	293
2.3.1.3. Genética de populações.....	297
2.3.2. Resultados.....	299
2.3.2.1. Morfometria geométrica.....	299
2.3.2.2. Genética de populações.....	301
2.3.2.3. Genética de populações do Morfotipo 01 (<i>Caranx crysos</i>)..	303
2.3.3. Discussão.....	305
2.4. CONCLUSÃO.....	308
3. BIOINVASÃO.....	310
3.1. INTRODUÇÃO.....	311
3.2. ISOGNOMON BICOLOR.....	314
3.2.1. Material e Métodos.....	317
3.2.1.1. Área de estudo.....	317
3.2.1.2. Eletroforese de aloenzimas.....	317
3.2.1.3. Análise estatística.....	318
3.2.2. Resultados.....	320
3.2.3. Discussão.....	324
3.3. LIMNOPERNA FORTUNEI.....	328
3.3.1. Material e Métodos.....	333
3.3.1.1. Área de estudo.....	333
3.3.1.2. Eletroforese de aloenzimas.....	334
3.3.1.3. Morfometria.....	334
3.3.1.4. Análise estatística.....	335
3.3.2. Resultados.....	337
3.3.2.1. Genética.....	337
3.3.2.2. Morfometria.....	343
3.3.4. Discussão.....	345
3.4. CONCLUSÃO.....	349
4. SAMBAQUIS E BIODIVERSIDADE DO HOLOCENO.....	351
4.1. INTRODUÇÃO.....	352
4.2. DADOS SECUNDÁRIOS: LAGOA DE ARARUAMA.....	355
4.2.1. Material e Métodos.....	358
4.2.2. Resultados.....	362
4.2.3. Discussão.....	368

4.3. DADOS PRIMÁRIOS: SAMBAQUI DA TARIOBA.....	373
4.3.1. Material e Métodos.....	375
4.3.2. Resultados.....	378
4.3.3. Discussão.....	381
4.4. CONCLUSÃO	382
5. SAMBAQUIS E PADRÕES DE BIODIVERSIDADE.....	384
5.1. INTRODUÇÃO.....	385
5.2. REVISÃO DA LITERATURA.....	389
5.2.1. Estabelecimento de Baselines.....	390
5.2.2. Bioinvasões.....	392
5.2.3. Padrões de Biodiversidade.....	394
5.3. SAMBAQUIS COMO PROXY DA BIODIVERSIDADE.....	395
5.3.1. Viés ou informação?.....	396
5.3.2. Origem dos dados e análises quantitativas.....	398
5.3.2.1. Sambaquis como amostradores da biodiversidade.....	400
5.3.2.2. Evolução nos Padrões de Biodiversidade.....	402
5.3.3. Resultados.....	405
5.3.3.1. Acumulações artificiais, mas nem tanto!.....	405
5.3.3.2. Mudança temporal dos padrões de biodiversidade.....	408
5.4. CONCLUSÃO	410
6. CONSIDERAÇÕES E CONCLUSÕES FINAIS.....	412
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	415

ÍNDICE VOLUME III

8. APÊNDICES.....	453
 8.1. GENÉTICA MARINHA.....	454
Duarte, M.R. & Silva, E.P. 2015. Apontamentos para uma história da Genética Marinha no Brasil. In: <i>Anais do V Congresso Brasileiro de Biologia Marinha</i> , Ipojuca, PE.....	455
 8.2. RECURSOS PESQUEIROS.....	459
Bonner, A.; Duarte, M.; Souza, R.; Monteiro-Neto, C. & Silva, E. Taxonomic status and evidence of natural selection in two morphotypes of <i>Coryphaena hippurus</i> Linnaeus 1758 (Perciformes: Coryphaenidae). Submetido à revista <i>Neotropical Ichthyology</i>	460
Duarte, M.R.; Tubino, R.A.; Monteiro-Neto, C.; Martins, R.R.; Vieira, F.C.; Andrade-Tubino, M. & Silva, E.P. Genetic and morphometric evidence that jacks (Carangidae) fished off the coast of Rio de Janeiro (Brazil) comprise four different species. Submetido à revista <i>Biochemical Systematics and Ecology</i>	482
 8.3. BIOINVASÃO.....	504
Duarte, M.R. & Silva, E.P. 2016. Invasões Biológicas. Pp. 27-79. In: <i>Biologia quase ao extremo</i> . Lima, N.R.W. (Org). Associação Brasileira de Diversidade e Inclusão, Rio de Janeiro.....	505
Lopes, R.P.; Duarte, M.R. & Silva, E.P. 2014. A Genética e as invasões biológicas: dois estudos de caso de bivalves invasores do Brasil. <i>Genética na Escola</i> 9:86-91.....	545
Lopes, R.P.; Duarte, M.R.; Campos-Villaça, R. & Silva, E.P. Estabilidade temporal dos níveis de variação gênica da espécie de bivalve invasor <i>Isognomon bicolor</i> (C.B. Adams, 1845) (Bivalvia, Isognomonidae) na Praia de Itaipu, Niterói – RJ. Aceito para publicação na <i>Revista Brasileira de Biociências</i>	559
Duarte, M.R.; Otegui, A.C.P.; Fernandes, F.C. & Silva, E.P. High levels of genetic and morphological variability in invasive <i>Limnoperna fortunei</i> (Dunker, 1857) populations in South America. Submetido à revista <i>Malacologia</i>	594
 8.4. SAMBAQUIS E BIODIVERSIDADE DO HOLOCENO.....	633
Beauclair, M.; Duarte, M.R. & Silva E.P. 2016. Sambaquis (shell mounds) and mollusk diversity in the past history of Araruama Lagoon, Rio de Janeiro, Brazil. <i>Pan-American Journal of Aquatic Sciences</i> 11(1):47-59	634
Rodrigues, F.B.; Duarte, M.R.; Souza, R.C.C.L., Soares-Gomes, A. & Silva. E. P. 2016. Holocene crustaceans from the Tarioba shell mound, Rio das Ostras, Rio de Janeiro, Brazil. <i>Check List</i> 12(2):1-5.....	648
 8.5. SAMBAQUIS E PADRÕES DE BIODIVERSIDADE.....	654
Silva, E.P.; Arruda, T.A.; Souza, R.C.C.L. & Duarte, M.R. Sambaquis e Biodiversidade Pré-histórica da Costa Brasileira. Aceito para publicação na Revista <i>Ciência Hoje</i>	655
Silva, E.P.; Pádua, S.C.; Souza, R.C.C.L. & Duarte, M.R. Shell Mounds of the Southeast Coast of Brazil: Recovering Information on Past Malacological Biodiversity. Aceito para publicação em <i>Zooarchaeology in the Neotropics: Environmental diversity and human-animal interactions</i> . Mondini, M.; Munoz, A.S. & Fernández, P.M. (Eds). Springer Editorial, Heidelberg, Germany	669
Souza, R.C.C.L; Lima, T.A.; Duarte, M.R. & Silva, E.P. Changes in patterns of biodiversity of marine mollusks along the Brazilian coast during the late Holocene inferred from shell-mound (sambaquis) data. <i>The Holocene</i> doi: 10.1177/0959683616645946.....	687

LISTA DE FIGURAS VOLUME I

Figura 1.1. Definição gráfica da Genética Marinha constituída na interseção entre a teoria da Genética de Populações, as técnicas de mensuração da variação gênica e a Biologia Marinha.....	37
Figura 1.2. Esquema representando as conclusões lógicas da teoria evolutiva darwiniana.....	43
Figura 1.3. Resumo esquemático das três observações e das duas deduções de Darwin expostas no capítulo 3 de <i>A Origem das Espécies</i> e que produzem uma definição mais objetiva da força de seleção natural. D= descendentes; P= parentais; N= tamanho da população e K= constante.....	44
Figura 1.4. Resumo esquemático das previsões teóricas da hipótese neutralista.....	53
Figura 1.5. Gráfico ilustrando o número de trabalhos publicados em revistas científicas com as palavras <i>Marine and Genetics</i> (base de dados: <i>Web of Knowledge</i> , acessado em dezembro de 2015).....	59
Figura 1.6. Gráfico ilustrando o número de trabalhos publicados em revistas científicas com as palavras <i>Marine and Genetics and Brazil</i> (base de dados: <i>Web of Knowledge</i> , acessado em dezembro de 2015).....	63
Figura 1.7. Os parâmetros estatísticos representados como uma distribuição espacial. A média fica no centro da distribuição em relação às medidas individuais das amostras. As amostras (observações) individuais se distribuem em torno da média, com uma distância que representa o desvio padrão de cada amostra em relação à média. O espaço composto por todas as observações realizadas (amostras), seu centro (média) e as distâncias (desvios) formam total de distribuição da média das medidas de cada amostra e das amostras individuais formam o espaço de variância.....	71
Figura 1.8. Histograma demonstrando seleção para tamanho de ostras. A- Distribuição dos fenótipos na população parental com média 10 cm, da qual foram selecionadas as ostras com tamanho superior a 15 cm para a reprodução. B- Distribuição dos fenótipos na população de descendentes das ostras maiores que 15 cm que apresenta, agora, média de 13 cm. Embora este exemplo seja imaginário, é sabido que distribuições como esta são obtidas em programas de melhoramento genético, especialmente se a seleção artificial é feita com populações heterogêneas. Nestes casos, o fenótipo médio da população muda, em poucas gerações, na direção determinada pela seleção.....	72
Figura 1.9. Medidas tomadas em estudos da forma da concha em mexilhões. A- Morfometria tradicional, na qual medidas lineares são obtidas e B- Morfometria geométrica, na qual as distâncias entre os pontos de contorno da concha são transferidas para um plano cartesiano.....	74
Figura 1.10. Diferentes formas de crescimentos em organismos. A- Crescimento Isométrico, no qual jovens e adultos da mesma espécie possuem a mesma forma, ou seja, todo o corpo do organismo cresce a uma taxa constante e B- Crescimento Alométrico, no qual jovens e adultos da mesma espécie possuem forma diferentes, ou seja, diferentes partes do corpo crescem a taxas diferentes.....	76
Figura 1.11. Resumo esquemático dos conceitos que servem para descrição da variação gênica em uma população (ver texto para explicação detalhada).....	80

Figura 1.12. Diferentes formas de cromossomos são definidas em função da posição do centrômero. Na forma metacêntrica, o centrômero divide o cromossomo em duas partes iguais; na sub-metacêntrica, o centrômero divide o cromossomo em duas partes desiguais; na acrocêntrica, o centrômero é sub terminal e na forma telocêntrica, o centrômero é terminal.....	82
Figura 1.13. Dogma Central da Biologia Molecular descrito pela primeira vez em 1958, por Francis Crick, um dos responsáveis pela proposição do famoso modelo de dupla hélice para o DNA. Em 1970 Crick voltou ao já popular Dogma em um artigo para <i>Nature</i>	89
Figura 1.14. Resumo do método de eletroforese de aloenzimas tendo, ao final, a interpretação mendeliana do padrão de bandas.....	89
Figura 1.15. Representação esquemática das três formas de DNA. A- genoma nuclear, nDNA, B- genoma mitocondrial, mtDNA (com os genes: COI, COII, ATP6, CytB, ND1, 16S e 12S) e C- unidade de repetição do rDNA de eucariontes (ETS: espaçador externo, 18S: região que codifica para rRNA 18S, ITS: espaçador interno, 25S/28S: região que codifica para rRNA 25/28S, NTS: espaçador externo não codificante ou espaçador intergênico). kb – kilobase. 1 kb é equivalente a 1000 pb (pares de bases).....	95
Figura 1.16. Representação esquemática da técnica de RFLP. Seguindo as setas: DNA extraído, adição das enzimas de restrição ao DNA extraído, eletroforese, coloração e padrão de bandas revelando o polimorfismo dos fragmentos de restrição interpretados no mapa de restrição ao lado. MM é o marcador molecular de tamanho. O alelo A é a sequência de DNA de 8 kb que não sofreu restrição, portanto, no gel é possível observar apenas uma banda no tamanho de 8 kb, como evidenciado pelo MM. O alelo B apresentou apenas um sítio de restrição, portanto, a sequência de 8 kb foi cortada em dois fragmentos (3,5 kb + 4,5 kb) como mostrado no gel, podendo seu tamanho ser conferido em relação ao MM. O alelo C, com dois sítios de restrição, originou os três fragmentos mostrados no gel (3,5 kb + 2,5 kb + 2,0 kb). Lembrando que kb = kilobase que equivale a 1000 pb (pb = pares de bases).....	99
Figura 1.17. Representação esquemática das etapas que constituem a técnica de PCR.....	101
Figura 1.18. Histograma de distribuição da frequências de cada um dos desvios da proporção esperada de cinco caras e cinco coroas em dez jogadas de uma moeda. A operação foi repetida 500 vezes. Os cinco desvios são: 1- quatro caras e seis coroas (ou vice-versa); 2- três caras e sete coroas (ou vice-versa); 3- duas caras e oito coroas (ou vice-versa); 4- uma cara e nove coroas (ou vice-versa) e 5- dez caras (ou 10 coroas).....	127
Figura 1.19. Representação dos haplótipos de quatro indivíduos de uma população natural de peixes. Observe que a sequência de nucleotídeos do indivíduo um é a mesma sequência de nucleotídeos do indivíduo três, isso significa que o haplótipo um é igual ao haplótipo três.....	130

Figura 1.20. Representação gráfica do F_{IS} que mede o índice de endocruzamento dentro de cada subpopulação. Ou seja, existe uma chance maior de reprodução entre os indivíduos dentro de cada subpopulação (A, B e C) do que entre os indivíduos da população A com aqueles das populações B e C; uma maior chance de cruzamento entre os indivíduos de dentro da população B do que com aqueles das populações A e C e assim por diante.....	137
Figura 1.21. Representação gráfica do F_{ST} que mede o índice de endocruzamento de cada subpopulação em relação à população total. Ou seja, quanto da variação total está entre as subpopulações e não dentro das subpopulações.....	137
Figura 1.22. Representação gráfica do F_{IT} que mede o índice de endocruzamento total que inclui tanto os efeitos do cruzamento preferencial dentro de cada subpopulação (F_{IS}), quanto o efeito da subdivisão populacional (F_{ST}).....	139
Figura 1.23. Representação do Modelo Continente-Ilhas de estrutura populacional.....	142
Figura 1.24. Representação do Modelo de Ilhas de estrutura populacional.....	143
Figura 1.25. Representação do Modelo <i>Stepping-Stone</i> de estrutura populacional.....	144
Figura 1.26. Há 50 milhões de anos atrás, as duas linhagens modernas tiveram um ancestral comum. Após 25 milhões de anos, as duas linhagens divergiram, cada uma delas tendo uma mudança de base. De modo que ficaram diferentes por duas bases. Após outros 25 milhões de anos, estas duas sequências acumularam mais duas bases de diferença (um nucleotídeo mudado para cada uma delas). Assim, após o total de 50 milhões de anos as duas linhagens divergiram em quatro bases (duas em cada linhagem). Segundo a hipótese do relógio molecular este processo de divergência é constante e regular e, no exemplo ilustrado, o seu tic-tac é de uma mudança de base a cada 25 milhões de anos, portanto, depois de 50 milhões de anos as duas sequências acumularam quatro diferenças de base entre si e, cada uma delas, duas diferenças em relação ao ancestral comum. O que significam 100 milhões de anos de divergência entre as sequências.....	153
Figura 1.27. Esquema representativo das rotas de ancestralidade de um conjunto de alelos. À medida que as gerações avançam no tempo (de cima para baixo), muitos alelos não deixam descendentes e, assim, se extinguem (círculos vazios). No processo ao contrário (de baixo para cima), a amostra de alelos observada na geração atual (círculos cheios) sofre uma série de eventos de coalescência nos quais os “n” alelos presentes na geração atual têm apenas “n – 1” alelos ancestrais até o ancestral comum mais recente.....	161
Figura 1.28. Árvores de espécies (contorno em traço contínuo) e árvores de genes (linha pontilhada). A- A topologia da árvore de espécies e da árvore de genes são iguais; B- A topologia da árvore de espécies e da árvore de genes são diferentes.....	165
Figura 1.29. Representação irônica das atitudes relativas ao processo de produção científica como descritas no texto.....	174
Figura 1.30. Representação esquemática das explicações para os altos níveis de variação gênica relacionadas às hipóteses neutralista e selecionista (N_e = tamanho efetivo de população; v = taxa de mutação e $H\uparrow$ = aumento da heterozigosidade). Maiores explicações sobre as hipóteses específicas podem ser encontradas no texto desta seção, como também em 4.2.1.3. <i>Controvérsias evolutivas</i> , especialmente a figura 1.4 no que diz respeito a hipótese neutralista.....	188

Figura 1.31. Representação de uma população composta por dois morfotipos de peixes (cauda branca e cauda preta) vivendo em simpatria. Estão exemplificados três <i>loci</i> usados para definir o status taxonomico dos morfotipos. O primeiro e o último <i>loci</i> não foram diagnósticos, uma vez que estão mostrando alelos partilhados pelos dois morfotipos, o que é uma indicação de ausência de isolamento reprodutivo entre eles. Contudo, o <i>locus</i> 2 é diagnóstico, já que indica que os morfotipos, mesmo vivendo em simpatria, apresentam isolamento reprodutivo, uma vez que não partilham alelos (nem um heterozigoto é observado).....	204
Figura 1.32. Gráfico de probabilidade da distribuição da Identidade Gênica de Nei para populações coespecíficas, congêneres e gêneros de uma mesma família.....	207
Figura 1.33. O enquadramento teórico da genética ecológica. São mostradas as fontes de variação gênica que são intrínsecas (mutação e recombinação) e extrínsecas (migração). Contudo, desta variação, aquela que interessa aos estudos de genética ecológica são aquelas que são sensíveis ao ambiente, ou seja, estão relacionadas com a adaptação. A seleção (natural) dos variantes populacionais herdáveis (variação gênica) é que vai determinar a adaptação.....	212
Figura 1.34. Esquema indicando algumas abordagens da Genética Marinha para os estudos sobre poluição do ambiente marinho.....	219
Figura 1.35. Esquema indicando a abordagem da Genética Marinha aplicada à pesca. Detalhes de estudos nessa perspectiva estão descritos no texto.....	223
Figura 1.36. Esquema indicando a utilização da Genética Marinha Marinha no desenvolvimento de estratégias para aquicultura.....	228
Figura 1.37. Algumas categorias definidas para as espécies exóticas, ilustrando gradiente crescente do potencial de invasão e a explicação causal da chegada, expansão ou extinção da espécie no novo ambiente.....	229
Figura 1.38. Representação gráfica de <i>bottlenecks</i> (efeito gargalo) na formação de populações invasoras.....	231
Figura 1.39. Estruturação taxonômica dos <i>taxa</i> de Pleurobranchomorpha. Os sete gêneros que foram amostrados no trabalho descrito são: <i>Pleurobranchaea</i> , <i>Euselenops</i> , <i>Pleurobranchus</i> , <i>Berthella</i> , <i>Berthellina</i> , <i>Tomthompsonia</i> e <i>Bathyberthella</i>	244
Figura 1.40. Tempos geológicos com indicação dos períodos de divergência e radiação de Pleurobranchinae.....	246

LISTA DE FIGURAS VOLUME II

Figura 2.1. Mapa da região de Cabo Frio, com os pontos de captura de <i>Coryphaena hippurus</i> destacados.....	270
Figura 2.2. Diagrama extraído de Thorpe (1983), demonstrando as probabilidades dos valores de identidade gênica (Nei, 1972) associados a populações da mesma espécie, espécies do mesmo gênero e espécies de diferentes gêneros. Os pontos em vermelho apontam as probabilidades associadas a identidade gênica encontrada entre os morfotipos estudados neste trabalho ($I = 0,935$).....	283
Figura 2.3. Dendograma gerado pela análise de aglomerados através da distância de Mahalanobis com os indivíduos analisados na PCA. Onde: D- dourado e P- palombeta.....	284
Figura 2.4. Mapa com os pontos de captura de xerelete.....	293
Figura 2.5. Vista lateral dos morfotipos de xerelete com a posição dos 18 marcos anatômicos: A - Morfotipo 01; B - Morfotipo 02; C - Morfotipo 03.....	296
Figura 2.6. Analise de componentes principais com os dados originais (A) e com os dados transformados por Procrustes (B).....	300
Figura 3.1. Ponto de coleta das amostras (Praia de Itaipu, Niterói – RJ).....	317
Figura 3.2. Dendrogramas construídos com base nas distâncias genéticas de Nei (1972) entre as amostras de <i>I. bicolor</i> analisadas. A- Usando o algoritmo das médias aritméticas não ponderadas (UPGMA) com os resultados de porcentagem de ocorrência em 1000 re-amostragens (<i>Bootstrap</i>) e B- Usando o algoritmo de agrupamento de vizinhos (<i>Neighbor-Joining-NJ</i>).....	323
Figura 3.3. Locais de amostragem de <i>L. fortunei</i> : Balneário Bagliardi (Argentina); Usina de Yacireta (Argentina, Paraguai); Usina de Itaipu (Argentina, Brasil, Paraguai); Corumbá, MS (Brasil) e Lago Guaíba, RS (Brasil).....	334
Figura 3.4. Medidas tomadas da valva esquerda de cada indivíduo de <i>L. fortunei</i> . (1) ls, (2) hei, (4) aam, (5) lar, (6) lig, (7) hp, (8) dpr, (9) pav, (10) pap, (11) pal.....	335
Figura 3.5. Dendrograma construído com base nas distâncias genéticas de Nei (1972) entre as populações de <i>L. fortunei</i> , usando o algoritmo das médias aritméticas não ponderadas (UPGMA) com os resultados de porcentagem de ocorrência em 1000 re-amostragens (<i>Bootstrap</i>).....	342
Figura 3.6. Plotagem de cinco populações contra os dois primeiros componentes principais usando os dados corrigidos para alometria com o método de Leonart e colaboradores (2000).....	344
Figura 4.1. Localização da Lagoa de Araruama e demarcação dos territórios de exploração e áreas (as formas ovóides delimitam os territórios de exploração e as zonas lagunares são diferenciadas por tons de cinza seguindo a mesma numeração dos territórios). Os círculos pretos representam os sítios arqueológicos, identificados por siglas (FO - Forte; BB - Boca da Barra; NA - Nacil; AI - Arco-íris, PE - Peró; PN - Salinas Peroano; ME - Meio; CE - Cemitério; IP - Ilha Palmer; AA - Abrigo Praia dos Anjos; PC - Ponta da Cabeça; CP - Colônia de Pescadores; MC - Morro da Concha; PS - Praia Seca).....	360

Figura 4.2. Número de espécies (riqueza de espécies) nos sítios arqueológicos e nos Territórios/Áreas da Lagoa de Araruama (as barras brancas representam a quantidade de gastrópodes e as barras pretas representam a quantidade de bivalves). Área 1: FO – Forte, BB – Boca da Barra, NA – Nacil, AI – Arco-íris, PE – Peró, PN – Salinas Peroano, ME – Meio; CE – Cemitério, IP – Ilha Palmer; Área 2: AA – Abrigo Praia dos Anjos, PC – Ponta da cabeça, CP – Colônia de Pescadores; Área 3: MC – Morro da Concha, PS – Praia Seca.....	362
Figura 4.3. Porcentagem de espécies por <i>habitat</i> (descritos na Tabela 4.1).....	364
Figura 4.4. Porcentagem de species por guildas alimentares.....	364
Figura 4.5. Espécies mais frequentes nos sambaquis: <i>Anomalocardia brasiliiana</i> (A), <i>Phacoides pectinatus</i> (B), <i>Anadara notabilis</i> (C), <i>Cerithium atratum</i> (D) and <i>Siratus senegalensis</i> (E). Fotos: Faria, R.G.....	365
Figura 4.6. Análise de cluster para os sítios arqueológicos.....	367
Figura 4.7. Análise de cluster para os territórios de exploração/áreas.....	367
Figura 4.8. Localização do Sambaqui da Tarioba.....	375
Figura 4.9. Estratigrafia do Sambaqui da Tarioba, mostrando as camadas culturais marcadas com fios de algodão cru. Foto: Souza, R.C.C.L.....	376
Figura 4.10. <i>Vouchers</i> das nove espécies de braquiúros identificadas no Sambaqui da Tarioba e depositadas na Coleção de subfósseis de sambaquis do Laboratório de Genética Marinha e Evolução-UFF. A: <i>Callinectes danae</i> ; B: <i>Callinectes sapidus</i> ; C: <i>Cardisoma guanhumi</i> ; D: <i>Goniopsis cruentata</i> ; E: <i>Menippe nodifrons</i> ; F: <i>Mithrax hispidus</i> ; G: <i>Ocypode quadrata</i> ; H: <i>Ucides cordatus</i> ; I: <i>Panopeus austrobesus</i> . Foto: Luz, Z.A.S.....	378
Figura 5.1. Em um sambaqui: A- Camadas estratigráficas evidenciadas na escavação; B- Detalhe; C e D- Triagem do material coletado; E- concha da espécie de gastrópode <i>Siratus senegalensis</i> e F- Concha da espécie de bivalve <i>Donax hilairea</i> . Fotos: Souza, R.C.C.L (A, B, C e D) e Faria, R.G. (E e F).....	388
Figura 5.2. Abordagem atualística utiliza vestígios arqueozoológicos para interpretar a paleoecologia com base nas características ecológicas das espécies do presente.....	391
Figura 5.3. Localização geográfica das três escalas estudadas: 1- Localização do Sambaqui da Tarioba; 2- Estado do Rio de Janeiro; 3- costa centro-sul do Brasil.....	399
Figura 5.4. Média e variação da distinção taxonômica para três locais. A: Saquarema-RJ; B: Cubatão-SP; Joinville-SC. Os triângulos representam os valores encontrados para os sambaquis e linhas (funil) representam o limite de 95%.....	409

LISTA DE QUADROS VOLUME I

Quadro 1.1. Descrição matemática da média.....	68
Quadro 1.2. Descrição matemática da variância.....	69
Quadro 1.3. Descrição matemática do desvio padrão.....	70
Quadro 1.4. Descrição matemática do cálculo das frequências genotípicas e gênicas..	119
Quadro 1.5. Descrição matemática da heterozigosidade por <i>locus</i>	120
Quadro 1.6. Descrição matemática da heterozigosidade esperada por <i>locus</i>	120
Quadro 1.7. Descrição matemática de nível de polimorfismo, número médio de alelos, número efetivo de alelos e heterozigosidade média.....	122
Quadro 1.8. Descrição matemática da Lei de Hardy-Weinberg.....	124
Quadro 1.9. Teste da hipótese de Hardy-Weinberg por qui-quadrado.....	129
Quadro 1.10. Número de sítios segregantes.....	131
Quadro 1.11. Diversidade haplotípica.....	132
Quadro 1.12. Diversidade nucleotídica.....	133
Quadro 1.13. Estimativa do F_{IS}	136
Quadro 1.14. Representação matemática do F_{ST}	138
Quadro 1.15. Representação matemática do F_{IT}	139
Quadro 1.16. Representação matemática da identidade gênica de Nei.....	147
Quadro 1.17. Representação matemática da distância gênica.....	148
Quadro 1.18. Representação de uma matriz de identidades e distâncias gênicas.....	148
Quadro 1.19. Representação matemática e gráfica da construção de um dendrograma de UPGMA.....	149
Quadro 1.20. Código genético com a apresentação da redundância de várias das trincas (conjunto de três bases).....	156
Quadro 1.21. Níveis de variação gênica em diferentes grupos taxonômicos.....	179
Quadro 1.22. As enzimas LAP e PGI.....	182
Quadro 1.23. Explicação neutralista para variação gênica e definição de tamanho efetivo de população.....	186
Quadro 1.24. Estimativa dos tamanhos de população para <i>Anomalocardia brasiliiana</i> e <i>Perna perna</i>	187
Quadro 1.25. Níveis de diferenciação geográfica em organismos marinhos.....	192
Quadro 1.26. Estruturação do mexilhão <i>Perna perna</i> na região de ressurgência em Arraial do Cabo (Rio de Janeiro, Brasil).....	194
Quadro 1.27. Diferenciação geográfica das frequências gênicas dos alelos do <i>locus Lap</i> em populações do molusco bivalve eurihalino <i>Anomalocardia brasiliiana</i> na costa brasileira.....	197

Quadro 1.28. Estudo comparativo (aloenzimas e DNA) sobre a estruturação genética das populações de <i>Mytilus edulis</i> na costa da Grã-Bretanha.....	201
Quadro 1.29. Estudo de sistemática bioquímica com a espécie de ouriço-do-mar <i>Paracentrotus gaimardi</i> no estado do Rio de Janeiro.....	205
Quadro 1.30. Previsões teóricas da hipótese neutralista sobre os valores de heterozigosidade em populações naturais.....	237

LISTA DE TABELAS VOLUME I

Tabela 1.1. Resumo dos critérios que devem ser usados na escolha de um DNA alvo, tipo de sequência e técnica para o trabalho com Genética Marinha.....	115
Tabela 1.2. <i>Softwares</i> , sua função, sítios onde estão disponíveis e referência original de apresentação do programa.....	170

LISTA DE TABELAS VOLUME II

Tabela 2.1. Sistemas de tampão utilizados.....	270
Tabela 2.2. Enzimas testadas, com abreviações, número de comissão (no E.C.) e número de subunidades.....	271
Tabela 2.3. Frequências alélicas para os morfotipos de <i>Coryphaena hippurus</i> . Abreviações: N: tamanho amostral; A, B, C: alelos.....	280
Tabela 2.4. Probabilidades (P) de equilíbrio de Hardy-Weinberg (H_o : união de gametas ao acaso).....	281
Tabela 2.5. Tamanho médio de amostra, número médio de alelos por <i>locus</i> , grau (%) de polimorfismo e heterozigosidades médias observadas e esperadas (Desvios-padrão entre parêntesis).....	281
Tabela 2.6. Heterozigosidades por <i>locus</i> (H_{OBS} = Heterozigozidade Observada; H_{ESP} = Heterozigozidade Esperada).....	282
Tabela 2.7. Índices de endocruzamento de Weir & Cockerham (1984) com os intervalos de confiança para as médias por <i>Bootstrap</i>	283
Tabela 2.8. Descrição dos marcos anatômicos utilizados no estudo com xerelete.....	295
Tabela 2.9. Frequencias alélicas para nove <i>loci</i> de xerelete. Abreviações: N: tamanho amostral; A, B, C: alelos; H_{OBS} : Heterozigosidade observada; H_{ESP} : Heterozigosidade esperada.....	302
Tabela 2.10. Frequências alélicas para quatorze <i>loci</i> em duas populações de <i>Caranx crysos</i> definidas pelo prorama Structure 2.3.3 (Pritchard et al. 2000). Abreviações: N: tamanho amostral; A, B, C: alelos; H_{OBS} : Heterozigosidade observada; H_{ESP} : Heterozigosidade esperada.....	304
Tabela 3.1. Frequências alélicas do bivalve <i>I. bicolor</i> para cada <i>locus</i>	321
Tabela 3.2. Níveis de variação gênica (Desvios-padrão entre parêntesis) para amostras de <i>I. bicolor</i>	321
Tabela 3.3. Probabilidades (P) de equilíbrio de Hardy-Weinberg (H_o : união de gametas ao acaso).....	322
Tabela 3.4. Índices de endocruzamento de Weir & Cockerham (1984) para <i>I. bicolor</i> .	322
Tabela 3.5. Matriz de similaridade genética. Acima da diagonal: identidade genética (NEI, 1978); abaixo da diagonal: valores de θ 2 x 2.....	322
Tabela 3.6. Frequências alélicas para nove <i>loci</i> gênicos de populações de <i>L. fortunei</i> da América do Sul. Abreviações: N = tamanho amostral; A,B,C = alelos; BB = Balneário Bagliardi; YA = Usina de Yacireta; IT = Usina de Itaipu; CO = Corumbá e LG = Lago Guaíba.....	339
Tabela 3.7. Níveis de variação gênica (Desvios-padrão entre parêntesis) para populações de <i>L. fortunei</i> . Abreviações dos nomes dos locais de coleta estão disponíveis na tabela 3.6.....	340
Tabela 3.8. Probabilidades (P) de equilíbrio de Hardy-Weinberg (H_o : união de gametas ao acaso). Abreviações dos nomes dos locais de coleta estão disponíveis na tabela 3.6.....	340

Tabela 3.9. Índices de endocruzamento de Weir & Cockerham (1984) para populações de <i>L. fortunei</i>	340
Tabela 3.10. Número de <i>loci</i> evidenciando deficiência/excesso de heterozigosidade sob o Modelo de Alelos Infinitos para detecção de <i>bottleneck</i> em populações de <i>L. fortunei</i> . Abreviações dos nomes dos locais de coleta estão disponíveis na tabela 3.6..	341
Tabela 3.11. Análise da variância molecular (AMOVA) para populações de <i>L. fortunei</i>	341
Tabela 3.12. Coeficientes de correlação para cada variável com os três primeiros componentes principais da variância total (dados logarítmicos e dados corrigidos pelo método de Lleonart e colaboradores 2000).....	343
Tabela 4.1. Descrição de categorias de <i>habitat</i>	360
Tabela 4.2. Lista de espécies de moluscos identificando os sítios onde foram encontradas, as guildas alimentares e os <i>habitat</i> (descritos na Tabela 4.1). Área 1: FO – Forte, BB – Boca da Barra, NA – Nacil, AI – Arco-íris, PE – Peró, PN – Salinas Peroano, ME – Meio; CE – Cemitério, IP – Ilha Palmer; Área 2: AA – Abrigo Praia dos Anjos, PC – Ponta da cabeça, CP – Colônia de Pescadores; Área 3: MC – Morro da Concha, PS – Praia Seca.....	361
Tabela 4.3. Os resultados dos testes de Kruskal-Wallis para riqueza de espécies obtidas da combinação aleatória de dois e três sambaquis da Área 1 <i>versus</i> riqueza de espécies obtidas de combinações aleatórias de dois e três sambaquis das Áreas 2 e 3 tomadas em conjunto.....	363
Tabela 4.4. Definição das espécies registradas nos sambaquis que podem possuir uso antropológico (1) e espécies que não tem uso conhecido (0), ocorrência relativa de espécies por sítio arqueológico e por território de exploração.....	366
Tabela 4.5. Riqueza de espécies relativa e datação para cada camada cultural.....	379
Tabela 4.6. Frequência de ocorrência (F) para cada espécie encontrada.....	379
Tabela 5.1. Sambaquis agrupados em 32 localidades com base nas semelhanças espaciais e ecológicos (os nomes dos Sambaquis estão disponíveis em Souza et al., 2011:28-43).....	402
Tabela 5.2. Diversidade taxonômica encontrada no Sambaqui da Tarioba (ST), costa do Rio de Janeiro (RJC) e da costa centro-sul do Brasil (CSC) do Brasil comparadas com a fauna de moluscos vivos para as mesmas áreas (RJC-LF e CSC-LF). A probabilidade de diversidade taxonômica ser diferente entre os grupos (sambaquis contra moluscos vivos) foi calculada com base em um teste de χ^2	405
Tabela 5.3. Resultados da analise de fidelidade quantitativa composicional para dados de sambaquis. F1 = % de espécies encontradas na assembleia viva que também são encontradas nos sambaquis, F2 = % de espécies encontradas em sambaquis que também são encontrados na assembleia viva.....	406
Tabela 5.4. Variância estandardizada da riqueza de espécies para três escalas: micro (sambaquis/pesquisas atuais), meso (localidades) e macro (estados).....	407

RESUMO

Nesta tese, aspectos relacionados à pesca, bioinvasão e evolução de padrões de biodiversidade foram abordados. Dois morfotipos da espécie *Coryphaena hippurus* Linnaeus, 1758 (dourado e palombeta) sofrendo, na costa do Rio de Janeiro, pressões de comércio e pesca diferenciadas foram investigados quanto ao seu *status taxonômico*. A análise dos dados indicou uma homogeneidade dos conjuntos gênicos dos dois morfotipos. Porém, foram encontradas diferenças significativas de variação gênica entre dourado e palombeta, o que sugere uma possível atuação de seleção natural sobre eles. Além de *C. hippurus*, indivíduos do xerelete, capturados ao longo da costa do Rio de Janeiro e desembarcados em Niterói, foram caracterizados por taxonomia clássica, morfometria multivariada e marcadores moleculares. O objetivo principal foi definir quantas e quais as espécies pescadas e produzir informações sobre a biologia e genética destas espécies. Foi possível identificar, pelo menos, quatro espécies sendo desembarcadas. Estes resultados são importantes para o desenvolvimento de programas de monitoramento e estratégias de gestão da pesca, além de fornecer informações sobre a ecologia e a evolução de Carangidae. Com relação ao problema da bioinvasão, a variação e estruturação genéticas foram investigadas para populações de moluscos bivalves a partir de uma perspectiva espacial (*Limnoperna fortunei*) e temporal (*Isognomon bicolor*) do problema. Os resultados indicaram níveis de variação gênica elevados e estáveis ao longo de oito anos (*Isognomon bicolor*) e uma prevalência de vetores não naturais mediando o processo de bioinvasão (*Limnoperna fortunei*). Por fim, o problema da biodiversidade foi abordado do ponto de vista da produção de inventários de referência baseados em dados de arqueozoologia de sambaquis. Como a diversidade biológica não é estática, compreender os seus padrões depende de mensurá-la tanto no presente, quanto seu comportamento em diferentes regiões e ao longo do tempo. Assim, a hipótese de evolução dos padrões de biodiversidade foi avaliada usando os inventários de referência produzidos, não sem antes testar o valor de dados e meta-dados oriundos de sítios arqueológicos que são reconhecidamente construções artificiais. Estes estudos incluíram inventários de metadados de moluscos para região da Lagoa de Araruama e dados primários de crustáceos do sítio arqueológico Sambaqui da Tarioba. Os testes referentes à quantidade e qualidade das informações sobre a biodiversidade que podem ser extraídas dos vestígios arqueozoológicos de sambaquis apresentaram resultados promissores.

ABSTRACT

In this thesis, aspects related to fishing, bioinvasion and evolution of patterns of biodiversity were addressed. Two morphotypes of the species *Coryphaena hippurus* Linnaeus, 1758 which are commercialized by different prices and suffer different fishing pressures in the coast of Rio de Janeiro had their taxonomic status investigated. Results indicated homogeneity of gene pools of the two morphotypes. However, significant differences in genetic variation were found between them, possibly due to natural selection. Added to *C. hippurus* study, individual of jacks fished along the coast of Rio de Janeiro and landed in Niterói, were characterized by classical taxonomy, multivariate morphometry and allozymes. The main objective was to define how many and which species are fished in the region in order to produce more precise information on the biology and genetics of these fisheries resource. It was possible to identify at least four species being landed, what have implications for monitoring and fisheries management, as well as to the enlight the ecology and evolution of Carangidae. Regarding the problem of bioinvasion, the genetic variation and population structure were investigated for populations of bivalves, both from a spatial (*Limnoperna fortunei*) and a temporal (*Isognomon bicolor*) perspective. The results indicated high and stable levels of genetic variation over eight years (*Isognomon bicolor*) and a prevalence of artificial vectors mediating the bioinvasion process (*Limnoperna fortunei*). Finally, the problem of biodiversity was approached from a historical standpoint taking archaeozoological vestiges to produce baselines. Furthermore, the hypothesis of evolution of biodiversity patterns was assessed using these data, not without testing the quality and quantity of information from data and meta-data from archaeological sites, which are well known as artificial buildings. Mollusc's meta-data were analysed for palaeoecological inferences related to Araruama Lagoon and primary data on crustaceans from Sambaqui Tarioba were used for producing baselines. Concerning the tests on the amount and quality of information on biodiversity that can be extracted from archaeozoological vestiges, results were promising.

INTRODUÇÃO GERAL

O primeiro capítulo desta tese (que constitui, também, o seu primeiro volume) foi escrito com o objetivo de apresentar uma introdução à teoria, às técnicas e aos problemas estudados em Genética Marinha. Nele são feitas uma tentativa de definição da “ciência” (na interseção dos vetores teoria da Genética de Populações, técnicas de mensuração da variação gênica e Biologia Marinha), bem como a construção de um histórico da disciplina no mundo e no Brasil.

O segundo volume desta tese apresenta um capítulo sobre genética marinha aplicada a problemas de pesca de dois recursos desembarcados no estado do Rio de Janeiro, um sobre genética aplicada a bioinvasão e dois sobre evolução dos padrões de biodiversidade. No Capítulo 02, foram produzidos dados de variação gênica, estrutura de populações e morfometria multivariada para o dourado (*Coryphaena hyppurus*) e o xerelete. No Capítulo 03, padrões de bioinvasão foram analisados no tempo (*Isognomon bicolor*) e no espaço (*Limnoperna fortunei*). No Capítulo 04, são apresentados inventários de referência de moluscos (dados secundários) e crustáceos (dados primários) obtidos a partir de sambaquis da costa do Rio de Janeiro. Neste capítulo são extraídas, ainda, inferências com respeito à evolução da biodiversidade destes *taxa* do Holoceno Recente ao presente (últimos 8000 anos A.P.). No Capítulo 05 é abordada a questão da quantidade e qualidade das informações sobre a biodiversidade que podem ser extraídas dos vestígios arqueozoológicos de sambaquis. Sambaquis são construções humanas realizadas na pré-história, desta forma, seus vestígios arqueozoológicos sofrem de viéses culturais e tafonômicos que, segundo alguns críticos, inviabilizam o uso das informações arqueozoológicas para inferências referentes à biodiversidade. O quinto capítulo se dedica a testar a eficiência dos dados de sambaquis na recuperação de padrões biodiversidade.

Ao longo de quatro anos (e meio) de doutorado os dados dos capítulos desta tese foram sujeitos a encontros, simpósios e congressos científicos; submissões a periódicos, algumas das quais lograram ser publicadas, aceitas ou estão, ainda, em processo de revisão pelos pares. O terceiro volume desta tese reúne todo este material, organizado por capítulo.

OBJETIVOS

O objetivo geral desta tese foi apresentar estudos nos quais a teoria e as técnicas da genética marinha e evolução foram aplicadas para solução de problemas relacionados à pesca, bioinvasão e padrões de biodiversidade, o que incluiu a análise de dados e meta-dados. Os objetivos específicos foram:

- 1) Apresentar uma introdução à teoria, às técnicas e aos problemas estudados em Genética Marinha, bem como as áreas nas quais ela pode ser aplicada colaborativamente;
- 2) Caracterizar genéticamente recursos pesqueiros (dourado e xerelete) desembarcados em portos do estado do Rio de Janeiro;
- 3) Utilizar a morfometria multivariada para discriminar possíveis espécies de Xerelete desembarcados no porto de Niterói, Rio de Janeiro;
- 4) Caracterizar variação e estruturação genética em espécies de moluscos bivalves invasores, tanto no tempo (*Isognomon bicolor*) e quanto no espaço (*Limnoperna fortunei*);
- 5) Avaliar a quantidade e qualidade das informações de vestígios arqueozoológicos de sambaquis sobre a biodiversidade do Holoceno;
- 5) Testar a eficiência dos vestígios arqueozoológicos de sambaquis na recuperação de padrões biodiversidade do Holoceno.

1. GENÉTICA MARINHA

1.1. INTRODUÇÃO

O que é Genética Marinha? A resposta para esta pergunta pode ser muito simples ou nem tanto. A Genética Marinha é a ciência que estuda a genética de populações de organismos marinhos. Contudo, ao defini-la assim, estão subentendidas outras definições. Por exemplo, a Genética de Populações, populações de organismos e organismos marinhos. Desta forma, é interessante começar por entender cada uma destas coisas.

1.1.1. Genética de Populações

A Genética é a ciência que estuda a transmissão das características biológicas (genótipo e fenótipo) de geração para geração. O genótipo pode ser definido como o conjunto de genes (unidade funcional da herança) que um organismo individual possui, enquanto o fenótipo é constituído pelas propriedades morfológicas, fisiológicas, bioquímicas, comportamentais e outras de um organismo, manifestadas ao longo da sua vida, que se desenvolvem pela ação de genes e sua interação com o ambiente. A genética pode ser estudada em nível de indivíduos e, também, em nível de populações (lewonin 1974; Futuyma 2009; Griffiths et al. 2006).

O termo “população” define-se como um grupo de indivíduos que acasalam uns com os outros, produzindo descendência. Uma espécie pode incluir uma ou mais populações diferentes, isoladas ou não. Uma população pode consistir em apenas alguns indivíduos ou em milhões deles, desde que esses indivíduos produzam, de fato, descendência. Um grupo de indivíduos que não pode se reproduzir não constitui uma população. Embora os indivíduos de uma população possam estar limitados a se reproduzir entre si devido ao isolamento físico, não existem impedimentos biológicos para sua reprodução com nenhum dos outros membros da sua espécie.

Dito isto, a Genética de Populações é a parte da Genética que estuda a proporção (ou frequência) dos diferentes genótipos e fenótipos nas populações e as mudanças nestas proporções ao longo das gerações. As mudanças de frequência dos genes ao longo das gerações resultam de seleção natural, mutação, migração e deriva genética. Ora, a mudança da composição genética das populações, ao longo das gerações, constitui aquilo que conhecemos como o processo evolutivo. Desta forma, estudar genética de populações é, também, estudar o processo evolutivo (Silva 2009b).

Assim, em sentido amplo, evolução pode ser definida como a mudança, ao longo do tempo, das proporções de organismos individuais que diferem geneticamente em uma ou mais de suas características. Em sentido estrito, tais mudanças expressam-se através da origem e subsequente alteração da frequência de alelos (variantes dos genes) ou genótipos de geração para geração dentro de populações (Hartl 2008).

Para que o processo evolutivo, em seu sentido estrito, ocorra, a primeira condição é que haja variantes dos genes (ou alelos) nas populações, ou seja, é necessária a existência de variação gênica. De outro modo, não seria possível que houvesse mudança nas populações ao longo das gerações. É possível concluir, então, que o trabalho de medir e caracterizar a variação gênica presente em populações naturais, bem como o entendimento dos mecanismos que determinam o seu padrão de distribuição nas populações, é o trabalho primário para se estudar a evolução no seu sentido estrito de Genética de Populações (Hartl & Clark 2010).

O estudo da Genética de Populações depende, portanto, de técnicas efetivas para observação e mensuração da composição genética presente nas populações naturais. A composição genética de uma população pode ser inferida através de estudos da variação morfológica dos organismos de uma população (como a morfometria tradicional e geométrica); do estudo dos cromossomos quanto a seus aspectos morfológicos e

funcionais, com a citogenética; de estudos da variação de proteínas, com a técnica de eletroforese de aloenzimas e, diretamente, a partir das variações encontradas no DNA (a base material dos genes).

1.1.2. Biologia Marinha

A Biologia Marinha é a ciência responsável pelo estudo dos seres vivos que têm como *habitat* os ecossistemas de água salgada, bem como as relações entre eles e os fatores abióticos. O ambiente marinho abriga um conjunto de formas de vida que representa a maior diversidade animal do planeta Terra (com exceção do grupo dos insetos que é o grupo de seres vivos com o maior número de espécies no planeta). A razão para maior representatividade da vida animal no mar se deve, possivelmente, ao fato da evolução ter se processado por mais tempo nos seus vários ambientes e, portanto, a sua diversidade de fatores ambientais (Soares-Gomes & Figueiredo 2009).

Os oceanos cobrem 362 000 000 Km² que representam, aproximadamente, 71% da superfície terrestre e possuem profundidade média em torno de 4 000 m, podendo atingir extremos de 11 000 m, aproximadamente, nas fossas abissais. No entanto, grande parte da vida marinha que se conhece, está concentrada em profundidades inferiores a 100 m (Soares-Gomes et al. 2009).

O ambiente marinho pode ser dividido em dois grandes domínios: o bentônico, que compreende a totalidade do substrato oceânico, e o pelágico, que corresponde a massa d'água total (Pereira & Soares-Gomes 2002; Bonecker et al. 2009; Paes & Monteiro-Neto 2009; Pereira & Soares-Gomes 2009). Nesses domínios encontramos as mais diferentes formas de vida, que, geralmente, são agrupadas em função do seu tamanho e hábito de vida. No domínio bentônico encontramos organismos que vivem no substrato, fixos ou não. No domínio pelágico encontramos organismos planctônicos,

que ficam à deriva dos movimentos oceânicos e organismos nectônicos, aqueles com boa capacidade natatória e, portanto, capazes de vencer os movimentos da massa d'água como as correntes, marés e ondas.

Diferenças de temperatura do ar entre os pólos e o equador dão origem a fortes ventos, que junto com a rotação da terra definem as correntes. As correntes profundas são definidas por variações da temperatura e salinidade que criam diferenças de densidade, mantendo uma circulação constante. As marés, produzidas pela força de atração da lua e do sol, são especialmente importantes nas zonas litorâneas, onde a vida marinha costuma ser bastante rica e variada. Divisões geomorfológicas também contribuem para formação de diferentes *habitat*, como as regiões oceânicas, com baixas concentrações de nutrientes e águas costeiras, dentro do limite da plataforma continental, onde as condições de nutrientes tornam a vida altamente favoráveis. Normalmente, também chamam as águas costeiras de zona nerítica, que é uma região que se encontra sob influência das marés. Nesta zona, encontram-se vários ecossistemas diferentes, como praias, estuários, manguezais, marismas, regiões de ressurgência e recifes.

O oceano é contínuo, sem separações definidas, sendo os gradientes de temperatura, salinidade e profundidade as principais barreiras para os organismos marinhos. Embora seja possível isolar cada um dos fatores ambientais e considerar a influencia de cada um sobre os organismos separadamente, o comportamento de um organismo em um dado momento será definido não por um simples fator exógeno, mas pela influência e interação de muitos fatores que atuam simultaneamente e pela ação de fatores endógenos. Assim, os organismos marinhos são funcionalmente adaptados a um conjunto de variáveis ambientais (Hellberg 2009).

1.1.3. Genética Marinha

Agora, talvez, a afirmação de que a Genética Marinha é a ciência que estuda a genética de populações de organismos marinhos possa ser menos hermética, contudo, melhor entendida na sua complexidade. Ou seja, estudar a Genética de Populações de organismos marinhos significa compreender a história evolutiva das espécies que vivem de forma variada em um ambiente multifacetado.

Mais que isso, estudar Genética Marinha significa, também, ser capaz de dominar e manipular conhecimentos referentes à Genética e Biologia Marinha e, também, as metodologias e técnicas associadas a cada uma delas. Contudo, não é saber “tudo de tudo”, mas ser capaz de criar um *locus* singular na interseção entre estes três vetores: a teoria da Genética de Populações, as técnicas de caracterização da composição genética dos organismos e a Biologia Marinha (Figura 1.1). É este lugar que constitui o poder (heurístico) e a fragilidade (epistemológica) desta disciplina. Contudo, é inegável que entre sua força e fraqueza, a aplicação da Genética Marinha tem ajudado a resolver problemas colocados pela investigação científica e nas aplicações das ciências do mar.

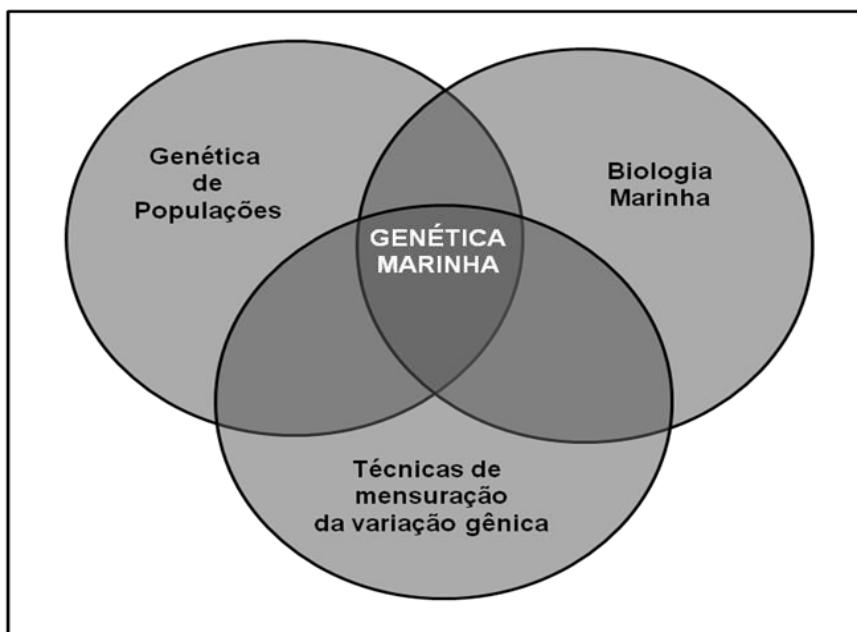


Figura 1.1. Definição gráfica da Genética Marinha constituída na interseção entre a teoria da Genética de Populações, as técnicas de mensuração da variação gênica e a Biologia Marinha.

1.2. HISTÓRICO

Antes de começar com uma pequena história de vida da Genética Marinha no mundo e, depois, sua chegada ao Brasil, é preciso falar da sua história familiar: Quem são o pai e a mãe desta disciplina? Ou seja, antes de falar dela mesma, é preciso discutir a sua filiação científica. A Genética Marinha, como é possível que já tenha sido intuído, se filia a teoria evolutiva darwiniana, mais especificamente, à teoria sintética da evolução e a sua dialética. É neste sentido que para bem entender o seu começo é preciso contar outra história, antes do seu começo.

1.2.1. Antes do Começo

A teoria da evolução biológica diz que a biodiversidade observada em todo o mundo, não esteve sempre aqui e, mais do que isso, que as espécies estão mudando ao longo do tempo. Em Biologia, portanto, evolução define uma forma de ver e entender a diversidade dos seres vivos, ou seja, é uma teoria da mudança. Contudo, como esta mudança é muito lenta e o tempo de que se fala está numa escala muito maior do que a que é possível perceber na vida cotidiana, é difícil, às vezes, imaginar como esse processo de mudança das espécies se dá. Essa dificuldade é muito comum.

É exatamente por conta da dificuldade de visualizar este processo de mudança das espécies no tempo que, antes que se pudesse entender de maneira adequada esse processo, era comum pensar que as espécies vistas hoje sempre existiram, no mesmo número e na mesma forma. A ideia de que as espécies não mudam é conhecida como fixismo. Essa ideia foi dominante durante muito tempo, porém, hoje em dia, diante de tantas evidências da mudança das espécies (domesticação de animais e plantas; fósseis de espécies extintas, incluindo espécies marinhas no alto do Andes; desenvolvimento de resistência a inseticidas em insetos e a antibióticos em bactérias; hibridização de espécies invasoras com espécies nativas etc.) só é possível manter este tipo de visão de

mundo se posicionando entre os extremos da ingenuidade e da má fé (Silva & Andrade 2012).

1.2.1.1. A teoria evolutiva darwiniana

A melhor compreensão do processo de mudança dos seres vivos veio com as ideias evolutivas de Darwin. Charles Robert Darwin, naturalista inglês, nascido na Inglaterra, que deu a volta ao mundo (1832-1837) em um navio, o *HMS Beagle* e que, por conta das suas muitas observações nessa viagem, produziu a mais importante teoria da evolução de que se tem notícia, descrita no seu livro, de 1859, *A origem das Espécies* (Darwin 1859). A teoria evolutiva darwiniana está entre as ideias mais importantes de toda a Biologia, por dois motivos, fundamentalmente. Primeiro, porque ela tem um caráter unificador, ou seja, ela, assim como a própria definição da vida e sua origem, a teoria celular e o conceito de gene integra todos os seres vivos como objeto de estudo único que a Biologia se propõe a entender. Segundo, porque a teoria evolutiva darwiniana é o “núcleo duro” de todas as teorias evolutivas modernas. Mas o que de tão importante Darwin escreveu no seu livro *A origem das Espécies*? O que nele ainda se mantém atual? Qual a novidade da teoria evolutiva darwiniana?

Afirma-se, geralmente, que Darwin criou a ideia de evolução, mas isto não foi criação dele. As ideias evolucionistas existiam desde o século XVII, sendo a teoria de Lamarck um belo exemplo disso (Rodrigues & Silva 2011; Arcanjo & Silva 2014; 2015). Outra afirmação comum, a respeito da teoria darwinista, é a de que a proposição do mecanismo de seleção natural seria sua grande novidade. Contudo, a tese da seleção natural, como mecanismo para evolução, já tinha encontrado outros defensores, como o próprio avô de Charles Darwin, Erasmus Darwin. Embora seja verdade que nos

trabalhos de Darwin o mecanismo de seleção natural apareça com maior importância e numa estrutura lógica nova, ainda assim o argumento não era novo.

A viagem no *HMS Beagle* e o acúmulo de dados que Darwin produziu para corroborar suas afirmações é outra novidade apontada para *A Origem das Espécies*, mas isto também não era novidade. O escocês Robert Chambers, contemporâneo de Darwin, já havia publicado o livro *Vestígios da história natural da criação*, em 1844, no qual reunia uma compilação imensa de dados para corroborar suas ideias evolutivas. Embora os dados de Chambers fossem de origem secundária, ou seja, compilados da literatura científica da época, a leitura do seu livro não deve nada, em termos de exemplos, àqueles presentes em *A Origem das Espécies* (Chambers 1844). Qual seria a novidade, então?

A grande revolução da teoria darwiniana foi a mudança na forma de encarar a variação presente entre indivíduos da mesma espécie. Até a emergência das ideias de Darwin, o pensamento reinante era tipológico e essencialista, ou seja, a identidade das espécies era fundada numa essência que, em última instância, definia a própria espécie. Dessa forma, as variações individuais eram tidas como desvio, como erro, em relação a um tipo padrão que a essência garantiria. Dito de outra forma, a essência de cada espécie garantia a todos os indivíduos uma homogeneidade, uma identidade em relação a este tipo padrão.

Até hoje, esta é uma prática estabelecida em taxonomia e sistemática que mantém nos museus o depósito dos tipos de todas as espécies descritas. Ora, vista assim, a variação observada era interpretada, consequentemente, como um desvio em relação ao tipo. Esta perspectiva tipológica e essencialista são marcadas pela noção de perfeição, que vem da ideia de criação.

Darwin, de forma diferente, encarava a variação individual sob uma perspectiva populacional. Para ele, a espécie não era um conjunto de indivíduos que deveriam se assemelhar a um tipo perfeito, mas um grupo (ou grupos) de indivíduos que partilhavam caracteres e tinham continuidade histórica através da reprodução. Essa nova forma de encarar a variação é uma revolução e está baseada numa perspectiva materialista da variação individual, que deixou de ser tida como estática, ou como o resultado da expressão imperfeita de uma essência imaterial, ou um ruído a ser evitado na atividade de ordenação (classificação) do mundo vivo, e passou a ser entendida como a realidade do mundo biológico e o material da evolução. A partir desta perspectiva materialista, Darwin pôde entender o processo de especiação como processo de conversão da variação entre indivíduos, dentro de uma determinada população, em variação entre populações diferentes, no tempo e no espaço. Esta é a segunda novidade da teoria darwinista: entender o processo de especiação como um processo de transformação de variação intrapopulacional em variação interpopulacional.

Essas duas novidades presentes no livro *A Origem das Espécies* têm consequências importantes, que foram percebidas imediatamente e causaram muita controvérsia. Primeiro, ficava estabelecido que a natureza das diferenças entre as espécies era a mesma das diferenças entre os indivíduos da mesma espécie. Essa interpretação era radicalmente contrária ao ponto de vista tipológico que encarava as diferenças entre as espécies como produto de variações em torno de uma essência de origem na Criação. Segundo, se o processo de formação de novas espécies dava-se pelo fracionamento da variação intrapopulacional em variação interpopulacional, a regressão desse processo nos levaria a conceber uma origem comum a todos os seres vivos, mas, isto também, se contrapunha violentamente à ideia de uma criação especial (Figura 1.2). Mais que isto, uma terceira conclusão: a evolução aconteceria, então, sem um propósito,

seria um processo de leis simples, para o qual não existia espaço para uma ideia de progresso (evolução era apenas mudança ao longo do tempo). Essas conclusões eram revolucionárias! (Solé-Cava et al. 2004a; 2004b; 2005)

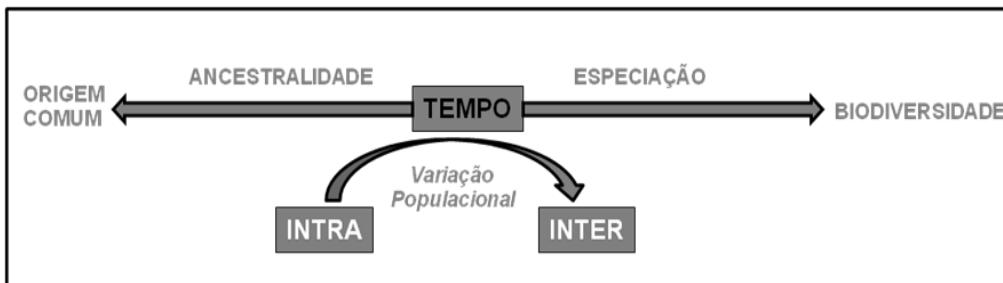


Figura 1.2. Esquema representando as conclusões lógicas da teoria evolutiva darwiniana.

No Capítulo 3 de *A Origem das Espécies*, denominado “Luta pela Existência”, Darwin apresentou três observações e duas deduções, que constituem uma nova roupagem para a velha ideia de seleção natural. Segundo ele, na natureza encontramos um número de parentais muito menor do que o de descendentes (primeira observação). No entanto, é fácil constatar que essa situação não ocorre de fato. Na realidade, o tamanho de populações naturais tem-se mantido mais ou menos constante ao longo do tempo (segunda observação). A dedução óbvia extraída dessas duas observações é a de que existe mortalidade de descendentes (primeira dedução). Nesse ponto, Darwin fornece sua terceira observação, que é, de fato, a grande novidade da sua teoria: existem diferenças entre os indivíduos de uma população, diferenças estas que podem aumentar ou diminuir as chances de o indivíduo ser bem sucedido no ambiente (terceira observação). Diante dessas três observações e de posse da primeira dedução, é possível entender que a mortalidade não ocorre ao acaso, mas em função das diferenças individuais (segunda dedução); ou seja, a mortalidade dos descendentes ocorre segundo

um processo de seleção que a natureza opera: uma seleção natural! Na figura 1.3 é apresentado um resumo das três observações e duas deduções de Darwin.

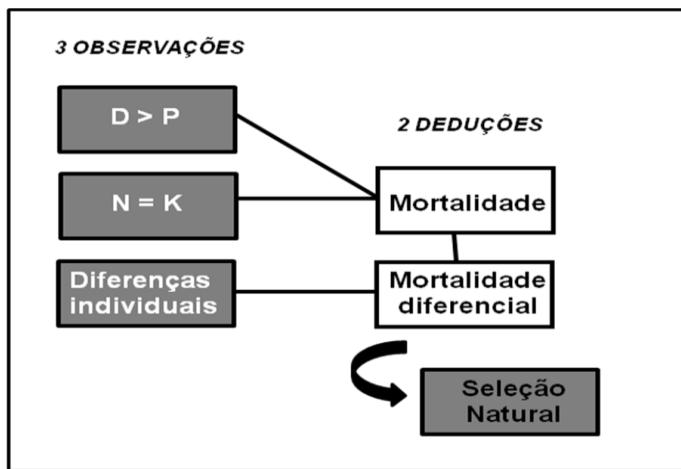


Figura 1.3. Resumo esquemático das três observações e das duas deduções de Darwin expostas no capítulo 3 de *A Origem das Espécies* e que produzem uma definição mais objetiva da força de seleção natural. D= descendentes; P= parentais; N= tamanho da população e K= constante.

Assim, a perspectiva materialista da variação se impunha, possibilitando uma interpretação extremamente elegante do mecanismo de Seleção Natural. A consequência de assumir um mecanismo como este, guiando a evolução, era estrondosa: um processo acéfalo, uma evolução sem desenho. Desse modo, tinha-se, naquele momento, uma definição do processo evolutivo que poderia ser resumida da seguinte forma: descendência com modificação guiada por força de seleção natural. Na seta do tempo, se seguisse para frente, haveria o processo de especiação e, consequentemente, a biodiversidade observada. Se, ao contrário, seguisse em direção ao passado, encontraria a descendência comum de todos os seres vivos.

A teoria darwinista da evolução era, portanto, revolucionária no seu tempo e permanece revolucionária ainda hoje. Ela traz uma interpretação completamente nova do mundo (a perspectiva materialista da variação) e possibilita o entendimento do processo de especiação e da natureza das espécies vivas (processo de transformação de variação intrapopulacional em interpopulacional), conferindo ao mecanismo de seleção

natural uma nova roupagem lógica. Mas nem tudo são flores e a teoria de Darwin tinha um problema.

1.2.1.2. A teoria sintética da evolução

Como já foi visto antes, para que um processo evolutivo ocorra, a primeira condição é que haja variação presente nas populações e que esta seja herdável. Darwin propunha que todos os organismos descenderiam de ancestrais comuns, através de um processo lento e contínuo de modificações, dirigido pela ação da seleção natural sobre os indivíduos. Porém, a teoria darwinista explicava a herança das modificações pelo processo da pangênese, no qual gêmulas, formadas em todas as partes do corpo, contribuiriam para as características adquiridas. De certa forma, tal teoria era uma atualização das ideias já formuladas por Lamarck, de herança dos caracteres adquiridos. Desse modo, a teoria darwinista não foi capaz, na sua época, de explicar nem a origem nem a natureza da variação, que era o material da evolução (Dobzhansky 1951; 1970; Freire-Maia 1988; Mayr 1993; Mayr & Provine 1998).

Foi a redescoberta dos trabalhos de Mendel, no início do século XX, que trouxe explicações novas sobre a herança que, daquele momento em diante, passou a ser definitivamente transferida dos pais para os filhos, através dos “fatores” hereditários. Mendel foi um dos primeiros a apresentar um modelo matemático preciso sobre um fenômeno biológico, seguindo as normas do método científico utilizado pela Física (Mendel 1865; 1869). Uma das novidades dos seus trabalhos residia no fato de se estudar a herança a partir de características discretas e pouco influenciadas pelo ambiente (Stern & Sherwood 1966; Charlesworth & Charlesworth 2009).

A genética mendeliana foi, inicialmente, encarada como um golpe fatal no darwinismo, pois, se o darwinismo tinha a preocupação de explicar a mudança

evolutiva, o mendelismo se preocupava com a estabilidade dos processos de herança, desprezava a variação contínua, base do darwinismo, e enfatizava a variação discreta.

Um dos principais problemas que contrapunha as ideias de Darwin e Mendel era que o modelo mendeliano de herança era baseado em fatores hereditários atomísticos, portanto, encarava o organismo como um conjunto de características discretas, herdadas de maneira invariante. Já a teoria evolutiva darwiniana, lidava com a variação contínua presente nas populações naturais, que era descrita pelos pesquisadores de história natural e estudada pelos biometristas. Em 1918, Ronald Fisher publicou um artigo (*The correlation between relatives on the supposition of Mendelian inheritance*) no qual demonstrava que todos os resultados obtidos pelos biometristas para herança de características de variação contínua poderiam ser derivados do modelo mendeliano, sendo necessário, para tanto, assumir a contribuição de vários *loci* e vários alelos com interação aditiva, nos casos mais simples. Este problema estava resolvido! (Fisher 1918)

Outro problema era que havia a época muito ceticismo em relação a capacidade da seleção natural de operar sobre populações mendelianas de modo a produzir os efeitos esperados pela teoria evolutiva darwiniana. Este problema foi resolvido com os trabalhos de Ronald Fisher, John Haldane e Sewall Wright. O trabalho destes três teóricos demonstrou matematicamente que era possível entender o processo evolutivo como um processo de mudança das frequências gênicas dentro das populações. As ideias de Fisher, Haldane e Wright, que possibilitaram a síntese, foram sumariadas e publicadas por volta de 1930. Fisher publicou o livro *The genetical theory of natural selection* em 1930 (Fisher 1930). Haldane também publicou um livro, *The causes of evolution*, em 1932, que contém uma série de palestras de divulgação da teoria evolutiva realizadas por ele e, ao final, um apêndice com o resumo de sua teoria matemática da

seleção natural (Haldane 1932). Wright, de modo diverso, publicou em 1931, um longo artigo, *Evolution in Mendelian populations*, na revista *Genetics* (Wright 1931).

Estava, assim, constituída a famosa “Teoria Sintética da Evolução-TSE”. A estrutura básica desta teoria é a de que a evolução é um fenômeno de duas faces: a produção de variação e a escolha de variantes. De maneira simplificada pode-se identificar a fonte da variação como sendo a mutação, que cria variação nova. A escolha dos variantes pode se dar tanto por seleção natural, que é uma força determinística, quanto ser estocástica (ou aleatória) pela ação da deriva genética (Shorrocks 1980; Silva 2001).

A seleção natural é um processo de mortalidade diferencial, ou seja, é absolutamente contingencial. Vai depender exclusivamente das condições presentes no momento e no local, bem como dos variantes presentes no grupo populacional. Desse modo, não é possível associar a este tipo de processo nenhum plano ou projeto. O processo de sobrevivência diferencial pode, algumas vezes, inclusive, criar circunstâncias problemáticas para algumas espécies.

A deriva genética, por outro lado, é a oscilação ao acaso da frequência dos variantes numa população, que vai ser sempre maior quanto menor for o tamanho da população. É como se fosse um jogo de moedas, você espera que as proporções de caras e coroas sejam 50%, contudo, com poucas jogadas, como jogando 10 vezes uma moeda para cima, você não se surpreende se caírem, por exemplo, 8 caras e 2 coroas. Se você jogasse a moeda para cima 1000 vezes você já ficaria surpreso se obtivesse 800 caras contra 200 coroas. Maior seria a sua surpresa ainda, com proporções semelhantes (8:2) em 10.000, 100.000 jogadas. Mas, em qualquer dos casos, você espera, sempre, um desvio das proporções esperadas, apenas pela ação do acaso. Na reprodução de uma população, de uma geração para outra, você vai ter, também, estes desvios das

proporções esperadas dos genótipos, apenas pela ação do acaso. É isto que chamamos de deriva genética.

E a mutação? A mutação é qualquer alteração ao acaso em um gene. Por serem ao acaso, poucas, muito poucas, são às vezes em que este processo produz um variante que seja mais eficiente do que o anterior. Assim, as mutações são, geralmente, deletérias. Raros são os eventos de mutações benéficas para uma população.

Dessa forma, podem-se tirar duas conclusões importantes a respeito do processo evolutivo descrito pela TSE. Primeira, que a evolução é um fato natural, inapelável, tanto quanto a gravidade. Isto porque, como a evolução, em última análise, resulta de mudanças nas frequências gênicas e a deriva genética, que é a ação do acaso, modifica as frequências gênicas de populações naturais de seres vivos, então, é impossível pensar em uma população que não esteja evoluindo. Não é necessário nem mesmo a seleção natural para que isto ocorra, embora a sua existência facilite e otimize o processo.

Segunda conclusão: como a evolução pode ser feita fundamentalmente baseada somente em forças como a ação do acaso (mutação, deriva) e da seleção natural, ou seja, atendendo a pressões imediatas do ambiente, o processo evolutivo não possui um planejamento. De fato, para gerar toda a biodiversidade observada hoje, o processo evolutivo dependeu de um período de tempo muito longo e muitas extinções. Logo, ideias como aquelas geralmente associadas ao processo evolutivo, como perfeição e progresso, não são adequadas. O processo evolutivo é um processo de alto custo!

Sabe-se que estas conclusões perturbam um pouco (talvez muito), mas é importante falar delas. Assim, para repetir algo que já foi dito neste capítulo algumas vezes até aqui, a TSE diz que a evolução é o processo de mudança da composição genética das populações ao longo das gerações. Este processo tem como resultado a ramificação das diferenças, seja entre indivíduos dentro da mesma população,

populações dentro da mesma espécie, espécies dentro do mesmo gênero e assim por diante. Dito de outra forma, a teoria evolutiva darwiniana descrita na figura 1.2, só que agora com uma explicação genética da variação. Contudo, em ciência, sempre que se acha uma resposta acaba-se encontrando outras perguntas...

1.2.1.3. Controvérsias evolutivas

Após o estabelecimento da Teoria Sintética, uma nova questão se impôs aos evolucionistas: qual a quantidade de variação gênica presente nas populações naturais? Duas respostas definiam duas escolas de pensamento, e uma controvérsia se estabelecia na TSE: a contradição entre Escola Clássica e do Balanço.

A Escola Clássica defendia que a quantidade de variação presente nas populações naturais deveria ser pequena, uma vez que a seleção natural atuante seria do tipo normalizadora, ou seja, uma seleção que tenderia a fixar nas populações os alelos mais bem adaptados, denominados alelos de tipo selvagem. A variação observada teria origem na mutação e seria de natureza transiente; estaria a caminho da fixação (caso fosse uma mutação benéfica) ou da extinção (caso fosse deletéria), sempre pela ação da seleção normalizadora.

A Escola do Balanço, por outro lado, argumentava que a heterozigosidade nas populações naturais deveria ser alta e mantida por seleção natural balanceada. Esse tipo de seleção inclui todas as formas de seleção natural que de alguma forma mantêm variação gênica nas populações, como, por exemplo, a vantagem dos heterozigotos e a seleção dependente de frequência. Para os balancistas, não poderiam existir alelos tipo (selvagem ou mutante), uma vez que a variação era estrutural.

A controvérsia entre a Escola Clássica e a do Balanço se estendeu da década de 1930 até a de 1960. Foi então que, em 1966, Richard C. Lewontin e John L. Hubby utilizaram a técnica de eletroforese de aloenzimas como método para estudar a quantidade de variação gênica presente nas populações naturais (Hubby & Lewontin 1966; Lewontin & Hubby 1966). Os resultados indicaram que os níveis de variação gênica presentes em populações naturais de *Drosophila pseudoobscura*, o organismo

com o qual foram realizados os trabalhos, eram muito altos. Cerca de 30% dos *loci* amostrados apresentavam polimorfismo (estavam segregando para mais de um alelo) para todas as populações estudadas. O número de alelos nesses *loci* variava de 2 a 6, com uma média de 3,3 alelos por *locus*. A proporção esperada de *loci* heterozigotos em um indivíduo era entre 8% e 15% dos *loci*. É preciso ressaltar que estes valores eram subestimativas da variação real, uma vez que a eletroforese de aloenzimas é um método conservador na amostragem da variação gênica (ver **1.3.5. Eletroforese de Aloenzimas**).

O quê, então, estes resultados indicavam? A conclusão era clara: a Escola do Balanço estava com a razão! Sim, mas não... Se os níveis de heterozigosidade observadas fossem mantidos por seleção natural balanceada, praticamente toda a população estaria mal adaptada, uma vez que cada indivíduo seria homozigoto pelo menos para um dos 8-15% dos *loci*. Isto é uma carga genética dura demais para qualquer população natural suportar. Assim, a Escola do Balanço ganhava, mas não levava. A contradição entre a Escola Clássica e do Balanço perdia, assim, todo o sentido. A pergunta agora era: que forças evolutivas mantêm os altos níveis de variação gênica nas populações naturais?

Em 1968, Motoo Kimura publicou um artigo científico na revista britânica *Nature*, no qual fez uma conta simples. Comparando estudos de moléculas de hemoglobina entre diferentes grupos de animais, ele percebeu que, durante a história evolutiva dos mamíferos, a taxa de substituição dos aminoácidos era muito baixa (Kimura 1968). Como a substituição de aminoácidos é o resultado da substituição de nucleotídeos em um códon, ele repetiu o cálculo, levando em consideração todo o genoma. Encontrou, então, uma taxa de substituição de nucleotídeos que era muito elevada. Segundo os cálculos de Kimura, um par de nucleotídeos estaria sendo substituído na população a cada dois anos em média, o que é uma evolução

extremamente rápida. Somado a isso, Kimura usou as estimativas de Lewontin e Hubby da proporção de *loci* heterozigotos encontrados para *D. pseudoobscura* e confirmou que, por seleção natural balanceada, nenhuma população seria capaz de manter os níveis de variação gênica observados com uma taxa de substituição tão elevada. A conclusão final de Kimura caiu como uma bomba: a variação gênica observada deveria, portanto, ser neutra.

As variantes alélicas observadas não deveriam ter valor seletivo algum. As forças evolutivas mais importantes passavam a ser, segundo esta perspectiva, a mutação e a deriva gênica. A questão evolutiva fundamental não era mais a sobrevivência do mais apto, mas a sobrevivência do mais sortudo! Estavam definidas, assim, duas novas perspectivas do processo evolutivo. De um lado, os selecionistas advogam que os níveis de variação gênica são mantidos por seleção natural (Lewontin & Krakauer 1973; Gillespie 1991; Karl & Avise 1992; McDonald 1994). De outro, os neutralistas defendem que a variação gênica é mantida por um equilíbrio entre taxa de mutação e tamanho efetivo de população (com o qual a força de deriva gênica está diretamente relacionada) (Skibinski et al. 1993; Andre et al. 2012).

A principal estratégia selecionista é tentar correlacionar as frequências de diferentes alelomorfos com fatores ambientais. Para isso, é necessário: 1– encontrar uma variação nas frequências gênicas ao longo do tempo que não possa ser explicada por mecanismos casuais; 2– correlacionar essa variação com algum fator ambiental; 3– observar que esse padrão se repete em outras unidades evolutivas. Mas só isso não basta. Para se aceitar que um polimorfismo seja mantido por seleção natural, é necessário, ainda: 4– que seja demonstrado que os alelos em questão produzem fenótipos diferentes; 5– que essas diferenças façam sentido no ambiente; 6– que esse sentido tenha um impacto sobre a adaptação. Estava definido aí um dos primeiros

programas de pesquisa da Genética Marinha, como será visto adiante neste capítulo (ver

1.6.4. Genética Ecológica).

A hipótese neutralista está baseada em um equilíbrio entre taxa de mutação e tamanho efetivo de população e faz previsões, *a priori*, a respeito do padrão de variação gênica (polimorfismo e heterozigosidade) que se espera encontrar nas populações naturais, em função do seu tamanho efetivo. Segundo os neutralistas, em uma população de tamanho efetivo finito, os alelos estarão entrando por mutação e se perdendo por ação da deriva. Como populações pequenas sofrem maior efeito da deriva, é esperado que, com uma taxa de mutação constante, as populações pequenas tenham menor nível de variação gênica (Figura 1.4).

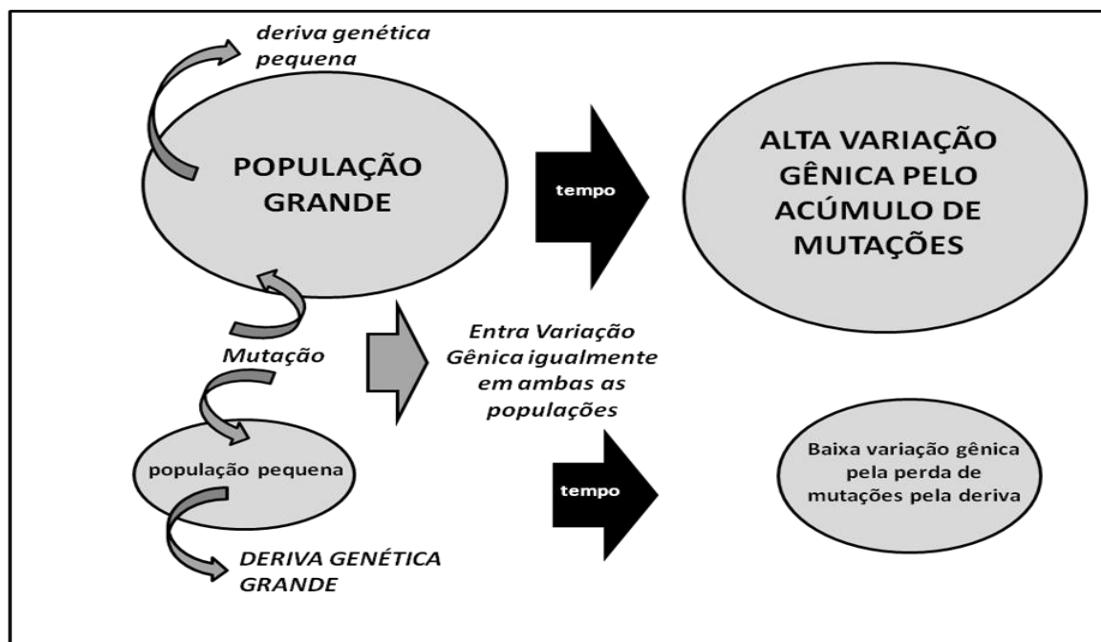


Figura 1.4. Resumo esquemático das previsões teóricas da hipótese neutralista.

As relações entre tamanho de população (deriva) e mutação são matematicamente definidas. Em função disso, uma das estratégias neutralistas é usar grandes bancos de dados que incluem informações sobre os níveis de variação gênica de vários grupos taxonômicos e comparar os resultados empíricos com aqueles definidos

matematicamente. Neste caso, também, a Genética Marinha entrou fortemente no debate, de modo que ainda é possível encontrar, hoje, geneticistas marinhos selecionistas, neutralistas e outros que, cansados desta controvérsia, seguiram outros caminhos (ver **1.6. Científico**).

O debate entre neutralistas e selecionistas está ainda presente na TSE e, também, na Genética Marinha. Contudo, mais recentemente esta controvérsia tem dado sinais de esgotamento, embora, não obrigatoriamente de solução. O modelo neutralista estrito não se manteve e foi, paulatinamente, sendo substituído pelo modelo quase neutro, que incorporou os dados que demonstram que as mutações são, geralmente, levemente deletérias (King & Jukes 1969; Otha & Gillespie 1996). Nenhum neutralista, por mais radical que seja, negará que algumas das mudanças evolutivas devem ser guiadas por seleção natural. Do mesmo modo, qualquer selecionista, hoje, tem de admitir a importância da deriva gênica no processo evolutivo. Se ainda não há solução no horizonte e os contendores mostrem sinais de cansaço, a controvérsia certamente deu bons frutos para o conhecimento da genética de populações no geral e de organismos marinhos no particular.

1.2.2. O Começo

A Genética Marinha é uma ciência jovem, contudo, que se expande rapidamente. Embora possam ser encontrados, na década de 1910, os trabalhos de Schmidt sobre variação quantitativa nos gêneros de peixe *Gadus* e *Zoarces* (*Racial Studies in Fishes* - Schmidt 1918; 1920); na década de 1920, o trabalho de Ford e Huxley (Ford & Huxley 1927) dedicado a estudar aspectos da genética do desenvolvimento de crustáceos anfípodes (*Mendelian inheritance and rates of development in Gammarus chevreuxi*) e, na década de 1950, trabalhos com grupos sanguíneos de atuns, salmões e bacalhau com

enfoque populacional (revisados no trabalho de W. de Ligny, *Serological and biochemical studies in fish populations* – Ligny 1969), o estudo da genética de organismos marinhos, utilizando as perspectivas populacional, evolutiva e ecológica (Genética Marinha), era restrito até o final da década de 60. Uma das razões para o desenvolvimento tardio de uma Genética Marinha foi, possivelmente, a dificuldade em manter cultivos em laboratório e realizar cruzamentos controlados com organismos marinhos. Outra possível dificuldade era o conhecimento limitado sobre aspectos ecológicos do ambiente marinho.

1.2.2.1. No mundo

É ao final da década de 1960 que a Genética Marinha se impõe como disciplina. Isto se dá devido ao desenvolvimento do método de eletroforese de aloenzimas para o estudo da variação gênica em populações naturais (lembre-se de: *1.2.1.3. Controvérsias evolutivas*). Assim, a Genética Marinha se estabelece no momento logo após a solução da controvérsia entre Escola Clássica e Escola do Balanço. Na seção 1.3, este método será descrito, no entanto, é importante, neste momento, fazer um pequeno parêntesis sobre o seu impacto para Genética de Populações e, também, seu papel no nascimento da Genética Marinha.

A utilização da eletroforese como técnica para separação de proteínas data da década de 50 (Smithies 1955; Hunter & Market 1957). Devido às diferenças de cargas elétricas, estrutura e tamanho as proteínas eram, rotineiramente, separadas e estudadas em laboratórios de Bioquímica. Contudo, como já se viu aqui (*1.2.1.3. Controvérsias evolutivas*) Dick Lewontin e Jack Hubby usaram a técnica para propósitos completamente distintos daqueles em que ela era usada até então. A grande novidade foi assumir o “Dogma Central da Biologia Molecular” como pressuposto (Crick 1958,

1970). Ou seja, se é verdade que o DNA é transcrito em RNA e, este, traduzido em proteínas, então, toda vez que forem observadas diferenças em proteínas, é possível inferir que elas se devem a diferenças nas sequências de DNA que as codificam: esta é a ideia do código genético! Desta forma, a técnica de eletroforese de proteínas passava a ser um método para estudar a variação gênica. Mas qual a diferença entre técnica e método?

Um método é uma técnica (eletroforese) a qual se associa um pressuposto teórico (Dogma Central da Biologia Molecular) de modo a que se possa responder uma pergunta (qual a quantidade de variação gênica presente nas populações naturais?). Desta forma, uma técnica de separação de proteínas da Bioquímica vai além dos seus horizontes e se torna um método para responder questões importantes da Genética de Populações. Pois bem, a partir do momento que a eletroforese de proteínas passou a ser empregada para o estudo da variação gênica de populações naturais, foi possível ter acesso a um grande número de *loci* em, virtualmente, qualquer organismo de qualquer ambiente, incluindo aí os, até então, pouco acessíveis organismos marinhos. A Genética Marinha vinha ao mundo para crescer!

Já em 1970, no V Simpósio Europeu sobre Biologia Marinha, em Veneza, havia um grande número de trabalhos dedicados a descrever a variação gênica em organismos marinhos (Battaglia 1972). Surgiu ali, então, a ideia de se fazer uma conferência exclusivamente dedicada aos problemas da recém-nascida Genética Marinha. Esta conferência ocorreu sete anos depois, também em Veneza, e resultou na publicação, em 1978, do livro *Marine Organisms: Genetics, Ecology and Evolution*, editado por Battaglia e Beardmore (Battaglia & Beardmore 1978). A Genética Marinha, nesta sua infância, tinha como preocupação central a produção de informações sobre os níveis de variação gênica em populações naturais de organismos marinhos, medidos através do

método de eletroforese de aloenzimas e a sua possível correlação com fatores ambientais. Discutia-se, ainda, neste momento, a possibilidade de se usar os dados assim produzidos em taxonomia.

Em 1983, no Reino Unido, o *Natural Environment Research Council* (NERC) resolveu estimular a Genética Marinha promovendo o seu dialogo com outras disciplinas da Biologia Marinha (tais como fisiologia, bioquímica e ecologia). Com este objetivo convidou Dick Koehn, referência em estudos de Genética utilizando organismos marinhos como modelos nos EUA, para uma conferência no XIX Simpósio Europeu sobre Biologia Marinha que foi realizado em Plymouth, em 1984. O interesse era estimular o contato e a discussão entre pesquisadores das duas áreas (Biologia Marinha e Genética) (Gibbs 1985). O resultado não foi dos mais excitantes, uma vez que os biólogos marinhos nos mostraram grande interesse pela abordagem. Contudo, o NERC persistiu no seu intento e estabeleceu um pequeno (e barato) programa para financiar projetos que promovessem a integração entre as duas áreas. Esta iniciativa inicial se mostrou extremamente profícua e determinou que, em 1989, fosse criado o programa *The Molecular and Genetic Analysis of Marine Systems* (GAMES). O GAMES foi encarado com cautela por pesquisadores externos ao Reino Unido. Primeiro porque se acreditava que, com relação aos sistemas marinhos, não havia uma comunidade de profissionais que pudesse se aliar com a comunidade de geneticistas. Segundo, havia certo ceticismo com relação à qualidade dos híbridos (Berry 1995).

Esta parte da história revela dois aspectos interessantes deste começo. Primeiro, que a naturalidade da Genética Marinha pode ser considerada europeia. Foi lá, na Europa, que a possibilidade de estabelecimento de uma Genética Marinha recebeu apoio institucional. Segundo, que nos EUA, neste momento, a Genética era um assunto de

geneticistas e Biologia Marinha de biólogos marinhos. Os híbridos existentes não eram estéreis, contudo, não reconheciam a subversão que se instalava.

No período em que o GAMES perdurou, se deu, em 1989, o XXIII Simpósio Europeu sobre Biologia Marinha, em Swansea-UK. Os trabalhos em Genética Marinha ali presentes evidenciaram a necessidade de realização de outra conferência inteiramente dedicada à disciplina da Genética Marinha (Ryland & Tyler 1989). Esta ocorreu, em 1992, em Bangor-UK e foi organizado pelo GAMES. Desta conferência resultou o livro *Genetics and Evolution of Aquatic Organism* editado por Beaumont e lançado em 1994, quinze anos depois do livro de Battaglia e Beardmore (Beaumont 1994). A Genética Marinha já era uma disciplina estabelecida, com aplicações em estudos de aquacultura, poluição e biologia da pesca. A tecnologia de DNA abria o seu caminho, principalmente aplicada a estudos filogenéticos com peixes.

A primeira conferência mundial nominalmente sobre Genética Marinha, o *International Workshop on Marine Genetics - Rio 1998* ocorreu no Brasil, em setembro, com a participação de vários daqueles que tinham feito a história da Genética Marinha. Neste encontro, ficou evidente o estabelecimento da tecnologia de DNA, pela primeira vez superando em número os trabalhos empregando o método de eletroforese de aloenzimas. A Genética Marinha estava preocupada, então, em estudos de biodiversidade molecular, conservação, poluição, aplicações em aquacultura, pesca, além de biotecnologia aplicada aos organismos marinhos. Os trabalhos dessa conferência foram publicados em um número especial da revista *Hydrobiologia* (Volume 420, Número 1) e, também, em forma de livro (Solé-Cava et al. 2000).

Assim, o estabelecimento da Genética Marinha ocorreu na década de 1990 e cresce continuamente desde então. Na figura 1.5 é possível visualizar o que foi dito aqui. É possível associar o crescimento da produção na disciplina à superação dos

obstáculos técnicos (estabelecimento do método de eletroforese de aloenzimas, o desenvolvimento da tecnologia do DNA recombinante). A partir do momento que os obstáculos são superados o crescimento é estimulado pela importância dos ecossistemas marinhos, tanto do ponto de vista econômico quanto ecológico. Além disso, a Genética Marinha é uma disciplina que integra diferentes áreas tendo, portanto, em certo sentido, um caráter interdisciplinar. A Genética Marinha hoje traz a possibilidade de aplicar a teoria e as técnicas disponíveis da genética aos os organismos marinhos, que têm muito para informar a respeito da evolução dos seres vivos.

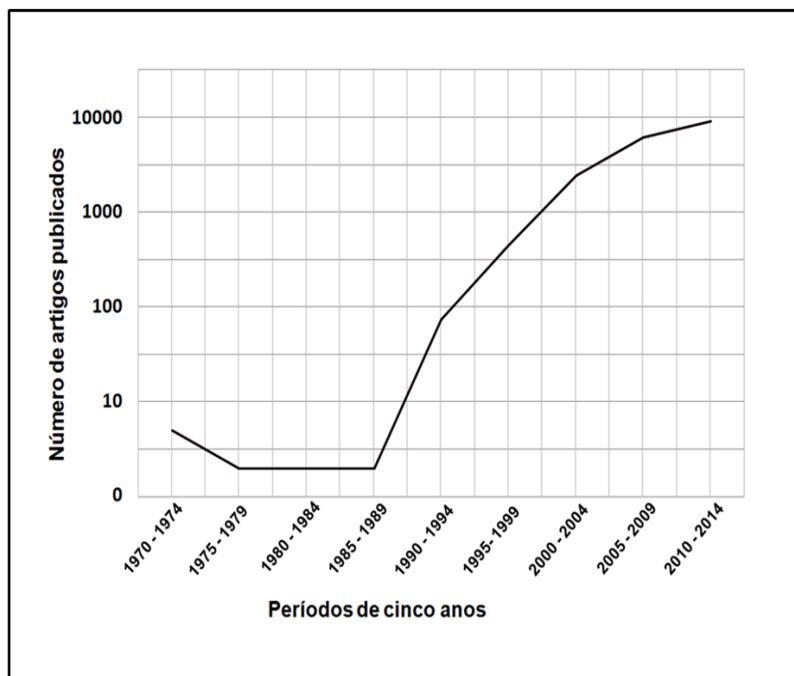


Figura 1.5. Gráfico ilustrando o número de trabalhos publicados em revistas científicas com as palavras *Marine and Genetics* (base de dados: *Web of Knowledge*, acessado em dezembro de 2015).

1.2.2.2. No Brasil

O início dos estudos de genética no Brasil data da década de 1930, quando estavam sendo estabelecidas as bases da teoria sintética da evolução. A genética era desenvolvida em institutos agronômicos como a Escola Agrícola Luiz de Queiroz

(Esalq) e o Instituto Agronômico de Campinas (IAC). Neles, a genética foi introduzida em pesquisas sobre o melhoramento de plantas. Na década posterior, foi a Universidade de São Paulo (USP) que ganhou proeminência com as iniciativas de André Dreyfus, pioneiro na divulgação da genética mendeliana e responsável pelo estabelecimento de um grupo de pesquisadores interessados na área (Salzano 1980; Souza et al. 2013). Entre eles, destaca-se Crodowaldo Pavan que, em 1944, defendeu sua tese de doutorado intitulada *Os peixes cegos das cavernas de Iporanga e a Evolução* que foi o primeiro trabalho de genética evolutiva animal feito no Brasil (Pavan 1944). É nesta época, também, em 1943, que a USP recebeu a visita de Theodosius Dobzhansky que influiu para que a década de 50 fosse caracterizada por intenso trabalho de genética e ecologia de drosofilas (mosca da fruta). O início da década de 1960 marcou o surgimento da genética humana como fruto da genética de drosofilas. A década de 1970 foi marcada pela expansão desta área, tanto em número de pessoas quanto de instituições.

É somente no início da década de 1980 que é publicado o primeiro trabalho em Genética utilizando um organismo marinho (*Isozymic differentiation of two sibling species of Squatina in South Brazil*). Este trabalho foi realizado por Antonio Mateo Solé-Cava na então Fundação Universidade do Rio Grande, hoje Universidade Federal do Rio Grande. Nele, dois tipos de tubarões que apresentavam diferenças morfológicas (morfotipos), mas eram considerados a mesma espécie, foram investigados, com o método de eletroforese de aloenzimas, para se saber se eram ou não a mesma espécie. Os resultados indicaram que os morfotipos eram, de fato, espécies diferentes e, portanto, não podiam ser pescados como uma espécie só. Ou seja, um trabalho clássico de aplicação da Genética Marinha para resolver problemas de sistemática e taxonomia, mas, também, com importantes consequências para o manejo da pesca e conservação das espécies (Solé-Cava et al. 1983).

É interessante notar algumas características importantes deste primeiro trabalho de Genética marinha no Brasil. Primeiro, ele surge à margem da história dos estudos de genética e evolução que se desenvolviam no Brasil e estavam centrados, principalmente, em instituições de pesquisa no estado de São Paulo. Em segundo lugar, ele se insere, metodologicamente, no programa de pesquisas sobre variação gênica em populações naturais inaugurado na década de 1960 por Lewontin e Hubby. Mais que isso, naquele momento, a utilidade do método de eletroforese de aloenzimas para estudos de sistemática e taxonomia estava se estabelecendo. Neste sentido, o trabalho com tubarões representa uma sintonia com a vanguarda da Genética Marinha no mundo. Por fim, ele apresenta uma ligação direta com as ciências do mar, seus problemas e interesses.

De fato, no seu primeiro momento, o crescimento e o estabelecimento da Genética Marinha brasileira estiveram mais ligados aos interesses imediatos das ciências do mar do que com a tradição de estudos em genética e evolução que se desenvolviam no país. Assim, foi mais a importância, econômica e ecológica, dos ecossistemas marinhos que motivou estes primeiros trabalhos. Só a partir da década de 1990 é que as atividades de vários grupos de pesquisa diretamente ligados a Genética (peixes de água doce, primatas etc.) começaram a re-orientar seu *expertise* metodológico-científico para o estudo de modelos marinhos. Esta re-orientação se deu, provavelmente, devido à solução de inúmeros problemas envolvidos com o estudo de organismos marinhos.

O trabalho com organismos marinhos apresenta alguns desafios para o pesquisador. Por exemplo, espécies marinhas são, geralmente, menos acessíveis do que as terrestres ou de água doce. Do mesmo modo, observações sobre comportamento e biologia destas espécies são mais difíceis. Por outro lado, os estudos de genética de espécies marinhas abrem oportunidades interessantes para as pesquisas em genética e

evolução. Por exemplo, muitos organismos marinhos apresentam padrões incomuns de dispersão, migração, bem como histórias de vida e fecundidade também singulares. Tudo isto traz efeitos para a estrutura genética das populações naturais que são importantes de serem compreendidos.

Outro fator importante para a re-orientação de grupos de pesquisa para Genética Marinha foi, certamente, o barateamento de muitas das técnicas de biologia molecular como a eletroforese de aloenzimas e, também, da tecnologia do DNA recombinante. A redução dos custos tornou possível a amostragem de um maior número de indivíduos para o estudo da genética das espécies. Por fim, o investimento do país em projetos ligados ao mar como, por exemplo, aqueles ligados ao mar territorial e a Zona Econômica Exclusiva (ZEE) e o interesse com a conservação e gestão da natureza são fatores que também podem ser apontados como tendo influência no estabelecimento da Genética Marinha no Brasil.

De um total de 27.523 grupos de pesquisa e 139.201 pesquisadores distribuídos pelas cinco regiões, existem hoje no Brasil três grupos de pesquisas que se denominam de Genética Marinha, todos situados no estado do Rio de Janeiro (Plataforma Lattes – <http://lattes.cnpq.br> e Diretório de Grupos de Pesquisa – <http://dgp.cnpq.br/diretorioc>). Contudo, se a produção dos grupos de pesquisa e pesquisadores é escrutinada, é possível encontrar um total de 403 grupos com produção científica relacionada à disciplina. Este total representa 1,5% dos grupos de pesquisa no país. Na figura 1.6 é possível observar a produção em Genética Marinha no Brasil desde a publicação do primeiro trabalho em 1983. A partir da década de 1990 o crescimento na área passou a ser intenso.

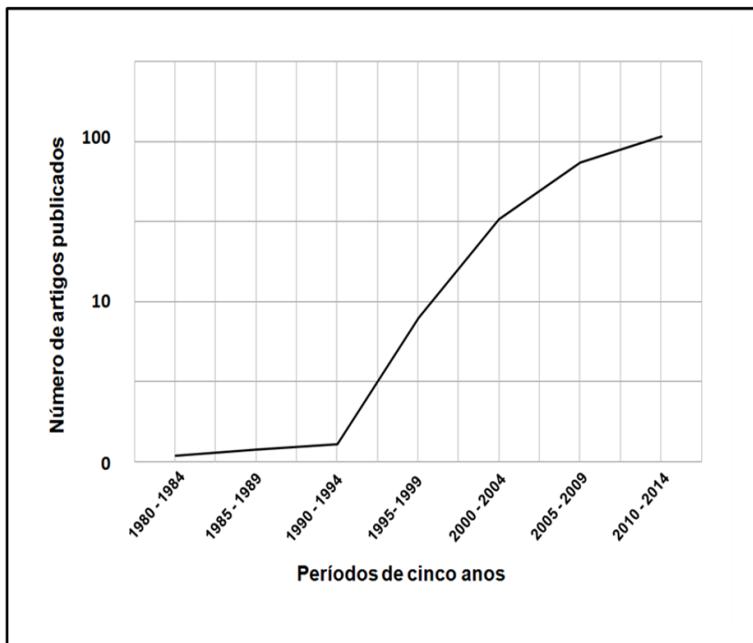


Figura 1.6. Gráfico ilustrando o número de trabalhos publicados em revistas científicas com as palavras *Marine and Genetics and Brazil* (base de dados: *Web of Knowledge*, acessado em dezembro de 2015).

1.2.2.3. Nos outros

A ciência é um empreendimento capitalista e, como tal, é lógico que se desenvolva preferencialmente nos países onde este modo de produção está estabelecido. A Genética Marinha é uma disciplina científica, assim, a sua história, como deve ter sido notado, se desenvolveu primeiramente no eixo Europa-EUA. Não é surpreendente, portanto, que “No mundo” seja um eufemismo para os países capitalistas desenvolvidos e exclua, por motivos óbvios, o resto do mundo. Contudo, a ubiquidade deste modo de produção vai incluindo, paulatinamente, os outros do mundo. Foi assim que o Brasil entrou nesta história.

É importante notar que desenvolvimento científico e desenvolvimento capitalista caminham juntos, portanto, a expressividade da Genética Marinha no Brasil pode ser entendida, também, como um aumento da expressividade científica do país como um todo que é resultado do aumento da importância econômica do Brasil em nível mundial

(e de alguns outros: Rússia, Índia, China e África do Sul-BRICS). Contudo, é necessário frisar que caminhar junto não é vir a reboque, o que significa que do mesmo modo que o modo capitalista de produção criou a ciência para seus objetivos de crescimento, a ciência faz o capitalismo crescer com o conhecimento que produz para manipular a natureza (Catani 1984).

Talvez este parêntesis, nesta pequena história da Genética Marinha, fosse desnecessário, mas o seu “(o) começo” em um “(no) mundo” restrito ao eixo Europa-EUA pode ter causado incomodo para alguns e merecia um esclarecimento, mesmo que breve. Mais que isso, o aparecimento do Brasil na linhagem histórica mundial da Genética Marinha, certamente tem muito de contingencial (história dos pesquisadores envolvidos), mas, também, não se pode negar, a uma casa, a olaria de onde seus tijolos vêm.

1.3. EMPÍRICO

Os problemas com os quais trabalha a Genética Marinha são, em última instância, problemas evolutivos. Utilizando modelos de estudo que são os organismos marinhos, o que traz problemas, mas, também, vantagens e perspectivas, já foi visto. Problemas evolutivos mesmo que aplicados a questões de interesse específico das ciências do mar, o que também já foi visto. Portanto, se os problemas em Genética Marinha têm uma natureza evolutiva e o material da evolução é a variação de caráter hereditário, então, a primeira etapa para se estudar Genética Marinha é, certamente, o desenvolvimento de técnicas e métodos de amostragem da variação. De fato, a própria origem da disciplina, como já foi discutida aqui, está ligada ao desenvolvimento de um método de amostragem de variação, a eletroforese de aloenzimas.

Técnicas e métodos fornecem a possibilidade de produção de observações e experiências que são utilizadas como evidências a favor ou contra as explicações que são dadas para os fenômenos da natureza. Quanto mais as explicações encontram apoio nas observações e experiências, maior é a sua base empírica, ou seja, mais observações e experiências existem que não contradizem a explicação.

Esta “crença” no poder das observações e da experiência para legitimar as explicações é uma forma de pensar o valor do conhecimento (seu conteúdo de verdade) que vem desde Aristóteles, na Grécia Antiga. Ele acreditava, em contraposição ao seu mestre Platão, que o conhecimento tinha uma natureza imanente (vinha do mundo concreto) ao invés de transcidente (com origem no mundo das ideias perfeitas). Contudo, foi no século XVII que o empirismo se estabeleceu com força, especialmente pelas ideias de Bacon, Locke e Hume. Estava em surgimento, naquele momento, o modo de produção capitalista e sua forma de conhecimento por excelência: a ciência moderna (Chalmers 1993).

O empirismo causou uma grande revolução, pois, graças à valorização do conhecimento científico, a sociedade passou a buscar resultados práticos, com o intuito de dominar a natureza e expandir o modo de produção capitalista. Desde o século XVII, a ciência e o capitalismo se expandiram juntos (como não poderia deixar ser) e muito. Contudo, a fé na “crença” empirista foi muito abalada, já que um detalhe importante foi percebido na própria prática científica: as técnicas (e os métodos mais ainda), só nos permitem observações e experiências que se encontram associadas a uma série de explicações sem base empírica, ou seja, os pressupostos (o que já foi visto na seção 4.2).

É neste sentido pouco “religioso” que o empírico será tratado aqui: técnicas e métodos que são usados para amostragem da variação de caráter hereditário que é o material da evolução e, portanto, da Genética Marinha.

1.3.1. Variação Fenotípica

Caráter é a denominação empregada para designar qualquer característica de um ser vivo que pode ser observada ou detectada e que, portanto, permite distinguir indivíduos de uma mesma espécie ou de espécies diferentes. Estes diferentes caracteres, que podem ser descritos para um determinado organismo, constituem o seu fenótipo. Alguns problemas estudados pela Genética Marinha focalizam a variação fenotípica. As características fenotípicas observadas em populações naturais podem ser de dois tipos: discreta ou contínua.

Diversos caracteres nos mais diferentes organismos podem ser divididos em classes fenotípicas distintas, neste tipo de variação o caráter apresenta apenas estados contrastantes. Na variação fenotípica discreta as classes são facilmente separadas uma das outras e os caracteres são chamados qualitativos. Como exemplo de caracteres qualitativos, pode-se citar a variação de cor nos espinhos de ouriços-do-mar (A espécie

Paracentrotus gaimardi apresenta cinco cores- preto, marron, cinza, verde e rosa). Entretanto, quando se observa os indivíduos nas populações naturais, pode-se identificar que, para muitos de seus caracteres, a variação fenotípica observada entre os indivíduos é contínua, ou seja, os fenótipos dos indivíduos não podem ser divididos em classes distintas, pois apresentam diferenças graduais. Na variação fenotípica contínua não há separação entre as classes, esses caracteres são chamados quantitativos e correspondem a grande parte das características. Como exemplo de caracteres quantitativos, pode-se citar a variação do peso em ostras e a variação no comprimento corporal de peixes.

O primeiro passo, para a análise de uma distribuição de variação fenotípica, é a estimativa da média aritmética. A média de uma amostra representa o centro da distribuição. Assim, quando se pensa na média de uma amostra deve-se lembrar de que este valor representa um valor central ao redor do qual todos os valores fenotípicos dos indivíduos da amostra estão distribuídos. A média é calculada através da soma de todas as medidas da amostra dividida pelo número de indivíduos analisados (Quadro 1.1).

Quadro 1.1. Descrição matemática da média.

A média pode ser descrita matematicamente como:

$$\bar{x} = \Sigma(x)/n$$

Onde:

\bar{x} - média;

Σ - somatório, ou seja, todos os valores devem ser somados;

x- cada amostra individual;

n- total de amostras (x) somadas.

Por exemplo, amostra-se o tamanho de cinco ostras com os valores em centímetros:

1, 2, 3, 4, 5

Então, a média será:

$$\bar{x} = (1+2+3+4+5)/5$$

$$\bar{x} = (15)/5$$

$$\bar{x} = 3$$

O segundo passo para a análise de uma distribuição fenotípica é a estimativa da variância, que nada mais é do que a dispersão dos valores fenotípicos dos indivíduos em torno da média. Assim, a variância da amostra é calculada como o desvio de cada uma das observações individuais em relação à média. Depois de calculado cada um dos desvios individuais, todos eles são somados e divididos pelo total de observações menos um. Um dado importante, embora mais técnico, é que, na verdade, o que é somado não são os desvios obtidos diretamente, mas os valores de cada um destes desvios ao quadrado (Quadro 1.2).

Quadro 1.2. Descrição matemática da variância.

A variância pode ser descrita matematicamente como:

$$s^2 = \Sigma(x - \bar{x})^2 / (n - 1)$$

Onde:

s^2 - variância;

\bar{x} - média;

Σ - somatório;

x - cada amostra individual;

n - total de amostras (x).

Por exemplo, no caso das cinco ostras amostradas, para as quais a média foi 3 ($\bar{x} = 3$), a variância, então, será:

$$s^2 = (1-3)^2 / (5-1) + (2-3)^2 / (5-1) + (3-3)^2 / (5-1) + (4-3)^2 / (5-1) + (5-3)^2 / (5-1)$$

$$s^2 = (-2)^2 / (4) + (-1)^2 / (4) + (0)^2 / (4) + (1)^2 / (4) + (2)^2 / (4)$$

$$s^2 = 4/4 + 1/4 + 0/4 + 1/4 + 4/4$$

$$s^2 = 1 + 0,25 + 0 + 0,25 + 1$$

$$s^2 = 2,5$$

Por fim, é interessante, também, se estimar o desvio padrão da média. O desvio padrão é uma medida importante por que informa sobre o quanto o conjunto de medidas individuais se distancia da média da população (Quadro 1.3). Assim se, por exemplo, duas gaivotas pescando, uma delas conseguiu capturar dez peixes e a outra nenhum, na

média, ambas estarão bem alimentadas, contudo, o desvio padrão será muito alto, ou seja, informa que não é bem assim, as duas observações (zero peixes e dez peixes) desviam muito da média que é cinco: uma observação é zero e a outra o dobro de cinco!

Quadro 1.3. Descrição matemática do desvio padrão.

O desvio padrão pode ser descrito matematicamente como:

$$s = \sqrt{s^2}$$

Onde:

s- desvio padrão;

s^2 - variância.

Por exemplo, para a variância da média igual de 2,5 ($s^2 = 2,5$) obtida no quadro 1.2, o desvio padrão será:

$$s = \sqrt{2,5}$$

$$s = 1,58$$

Na figura 1.7 tudo que foi dito até aqui sobre a média, a variância e o desvio é apresentado de maneira visual, de maneira que, os mesmos conceitos, possam ser apreciados na forma verbal, matemática e gráfica. Cada qual com a sua importância e função.

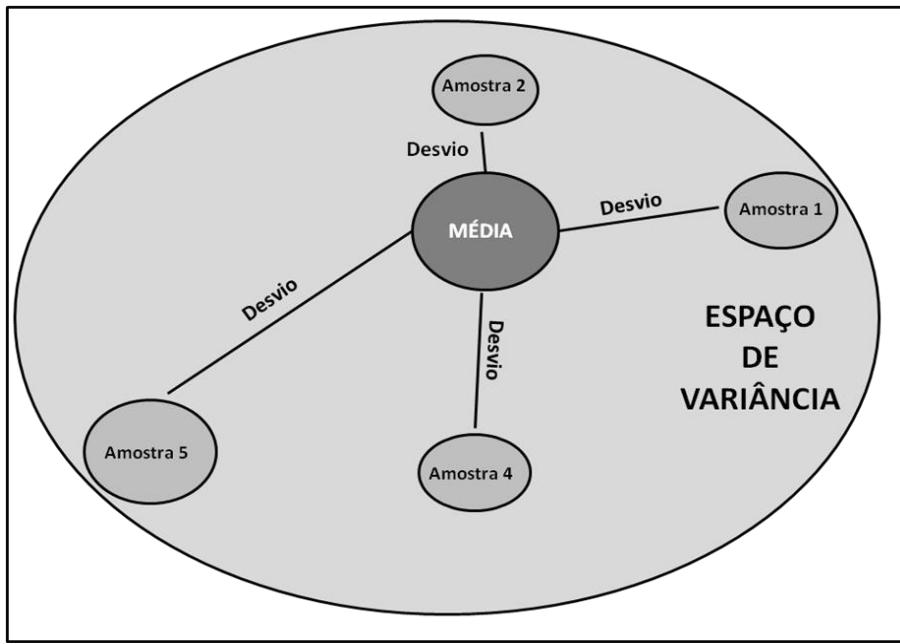


Figura 1.7. Os parâmetros estatísticos representados como uma distribuição espacial. A média fica no centro da distribuição em relação às medidas individuais das amostras. As amostras (observações) individuais se distribuem em torno da média, com uma distância que representa o desvio padrão de cada amostra em relação à média. O espaço composto por todas as observações realizadas (amostras), seu centro (média) e as distâncias (desvios) formam total de distribuição da média das medidas de cada amostra e das amostras individuais formam o espaço de variância.

Para o estudo e aplicação da Genética Marinha estas medidas são extremamente relevantes. Por exemplo, se em um cultivo de ostras deseja-se obter uma população com indivíduos maiores, faz-se uma seleção artificial escolhendo-se para reprodução os maiores indivíduos desta população. Contudo, a avaliação dos resultados do processo de seleção artificial depende da possibilidade de se descrever as populações de ostras (antes e depois da seleção artificial). Isto é feito com base nos valores de média, variância e desvio padrão dos valores fenotípicos individuais destas populações (Figura 1.8).

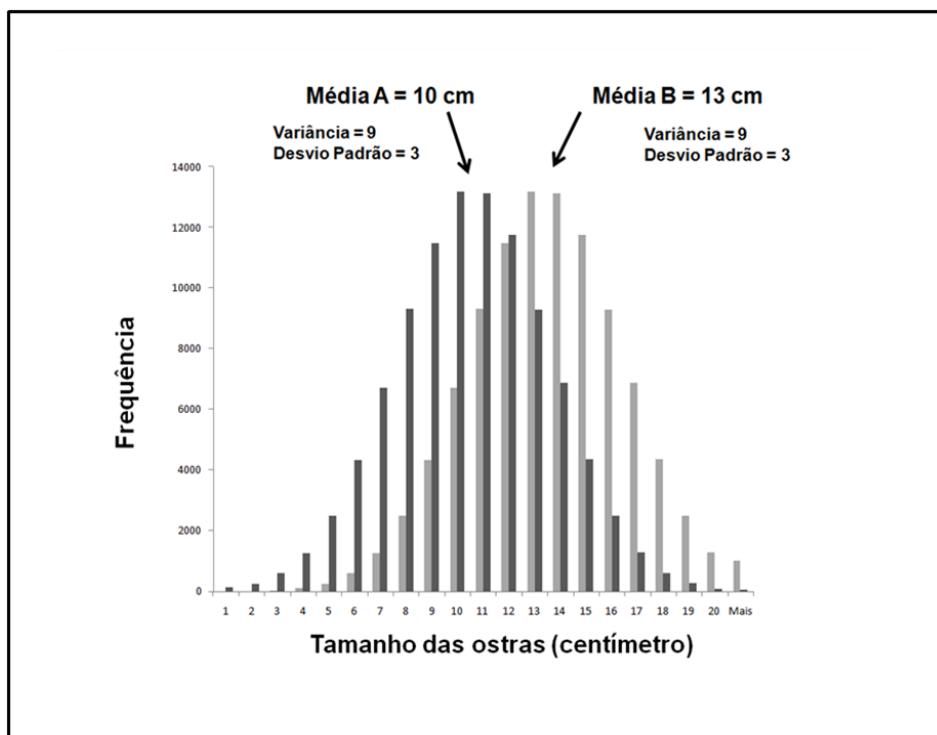


Figura 1.8. Histograma demonstrando seleção para tamanho de ostras. A- Distribuição dos fenótipos na população parental com média 10 cm, da qual foram selecionadas as ostras com tamanho superior a 15 cm para a reprodução. B- Distribuição dos fenótipos na população de descendentes das ostras maiores que 15 cm que apresenta, agora, média de 13 cm. Embora este exemplo seja imaginário, é sabido que distribuições como esta são obtidas em programas de melhoramento genético, especialmente se a seleção artificial é feita com populações heterogêneas. Nestes casos, o fenótipo médio da população muda, em poucas gerações, na direção determinada pela seleção.

A forma foi a primeiro tipo de variação fenotípica que o Homem utilizou para identificar os organismos. Desde que a ciência vem buscando métodos para uma classificação dos seres vivos, o principal critério para a identificação das espécies é seu conjunto de caracteres morfológicos - a sua forma, em última análise. Do ponto de vista da Genética Marinha, a existência de variações de forma entre os organismos, tanto dentro das populações quanto entre populações e espécies, é extremamente importante.

1.3.2. Morfometria

Morfometria é o estudo matemático das formas de objetos pertencentes à mesma população. Uma das suas aplicações é na identificação e na caracterização de populações que podem assumir formas ou tamanhos diferentes conforme o genótipo e o ambiente em que se desenvolvem (Monteiro & Reis 1999.). Sendo assim, a Morfometria é uma ferramenta muito importante na Genética Marinha.

Existem duas técnicas para definir, matematicamente, a forma dos organismos. A maneira mais simples, chamada morfometria clássica ou tradicional, é caracterizada pelo estudo da variância e covariância de medidas de distância entre pares de pontos entre estruturas. A morfometria tradicional descreve como e quanto estas medidas variam, bem como o quanto elas estão relacionadas entre si. As medidas usadas em morfometria tradicional podem ser obtidas tanto com uma régua (no caso de medidas como largura, comprimento, altura) como com um paquímetro (no caso de muitas medidas diferentes como distâncias entre diversos pontos definidos no corpo do organismo).

Inicialmente, os estudos da variação da forma eram feitas através de comparações de várias medidas, uma a uma. No entanto, devido ao fato de que os organismos são multidimensionais, notou-se que, mesmo se tomando várias medidas, isto não era suficiente para descrever a forma como um todo. A partir deste momento começaram a surgir os métodos multivariados usados atualmente. Neste caso, leva-se em consideração, simultaneamente, os diferentes níveis de variância e covariância entre as medidas. Devido à crescente facilidade de acesso às técnicas computacionais, as técnicas multivariadas tradicionais têm sido amplamente usadas, como Análise de Componentes Principais, Regressões Múltiplas, Análises Discriminantes entre outras.

Outro método utilizado na análise estatística das formas é a morfometria geométrica. Esta é uma ferramenta muito útil para detectar variações morfológicas indetectáveis por estudos de morfometria tradicional. A análise realizada por técnicas de morfometria geométrica está baseada na teoria de coordenadas cartesianas. No caso específico da morfometria geométrica, estes pontos cartesianos são denominados marcos anatômicos (ou *landmarks*, no inglês). As coordenadas informam sobre as posições relativas dos marcos anatômicos e, deste modo, permitem a reconstrução da forma estudada. Esta metodologia envolve uma série de técnicas que visam descrever e representar a geometria das formas estudadas. Na figura 1.9, o tipo de medida tomada para as análises de morfometria tradicional e geométrica estão representados para uma concha de mexilhão.

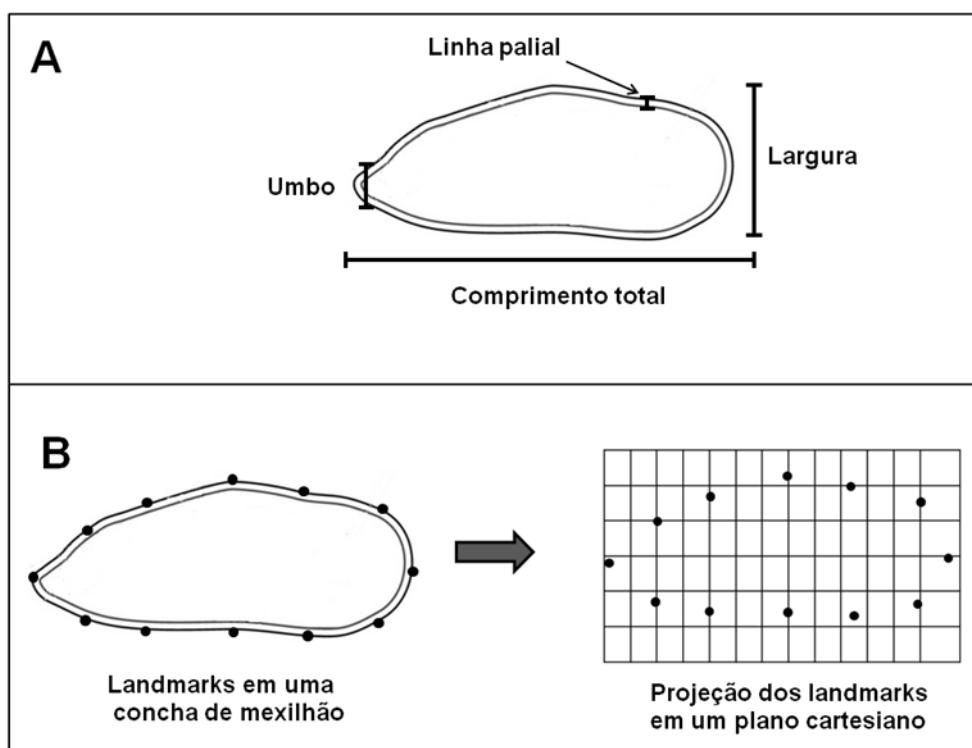


Figura 1.9. Medidas tomadas em estudos da forma da concha em mexilhões. A- Morfometria tradicional, na qual medidas lineares são obtidas e B- Morfometria geométrica, na qual as distâncias entre os pontos de contorno da concha são transferidas para um plano cartesiano.

Tanto a morfometria clássica quanto a geométrica são capazes de definir de maneira eficiente o padrão de forma de populações de organismos, uma vez que o seu crescimento seja isométrico. Ou seja, se o desenvolvimento ontogenético (da célula ovo, passando pelo embrião e até a fase adulta) mantiver um padrão de crescimento no qual as diversas partes do corpo do organismo crescem simultaneamente. Dito de outra forma, a medida de cada parte do corpo é uma função linear da outra. Assim, para comparar um organismo grande com um pequeno, basta a simples divisão por um fator de escala, por exemplo, o comprimento ou a altura. Nesta situação, como a forma é totalmente independente do tamanho, os organismos grandes e pequenos apresentam a mesma forma, sendo, os pequenos, miniaturas dos grandes.

O problema da morfometria, tanto tradicional quanto geométrica, está na situação em que diferentes partes do corpo crescem a taxas diferentes. Ou seja, à medida que o tamanho aumenta, a forma do corpo muda e a relação entre as medidas das diferentes partes do corpo não é mais linear. Neste caso, organismos grandes e pequenos terão, obrigatoriamente, formas diferentes devido ao crescimento diferencial das diversas partes do corpo. A consequência desta peculiaridade do desenvolvimento é que não há mais um fator linear de escala. Este crescimento, em que há uma mudança de forma acompanhando a mudança de tamanho, é denominado de crescimento alométrico.

Em análises morfométricas, o foco é a comparação entre a forma dos organismos. Neste contexto, a alometria é um grande problema, uma vez que as mudanças nas dimensões relativas de partes do corpo estão correlacionadas às mudanças no tamanho total. Nestes casos, grande parte da variação dos dados morfométricos deve-se a diferenças no tamanho dos organismos que são, geralmente,

consideradas como fontes contingenciais de variação. A figura 1.10 exemplifica o crescimento isométrico e alométrico.

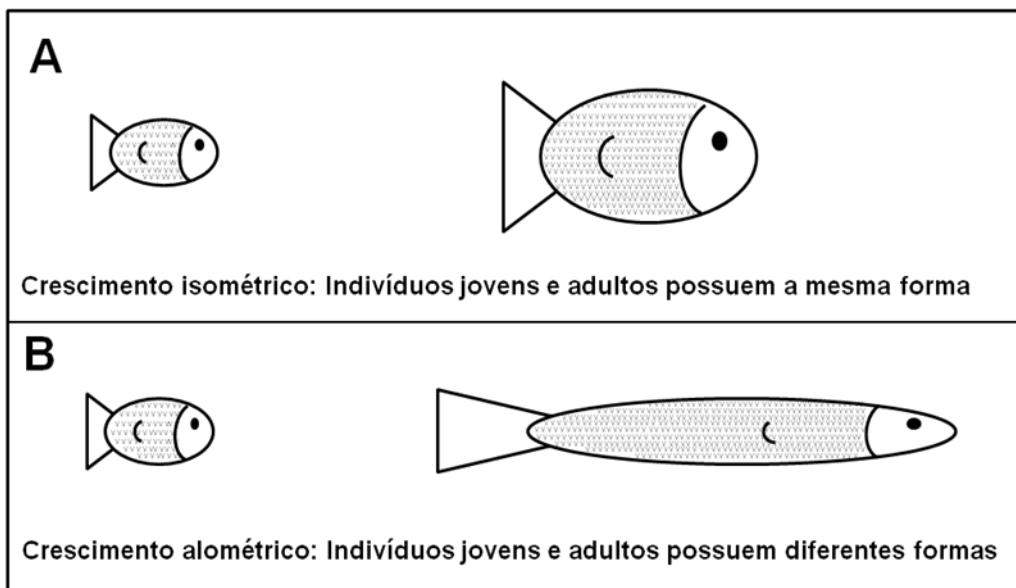


Figura 1.10. Diferentes formas de crescimentos em organismos. A- Crescimento Isométrico, no qual jovens e adultos da mesma espécie possuem a mesma forma, ou seja, todo o corpo do organismo cresce a uma taxa constante e B- Crescimento Alométrico, no qual jovens e adultos da mesma espécie possuem forma diferentes, ou seja, diferentes partes do corpo crescem a taxas diferentes.

Existem vários procedimentos estatísticos que têm o intuito de eliminar a influência do tamanho nas medidas de variação morfológica. O procedimento mais simples consiste em obter a razão entre cada medida e outra considerada como um estimador do tamanho, geralmente o comprimento total. Este procedimento realmente reduz os dados à mesma magnitude, porém não remove a influência não-linear do crescimento.

A despeito destes problemas, a variação da forma, sob a ótica da morfometria, tem sido descrita para espécies como peixes-rei do gênero *Odontesthes*, para o gastrópodo *Littorina saxatilis*, lagostas da espécie *Homarus americanus* e estudos taxonômicos em diversas espécies de mexilhões. Na seção **1.6. Científico**, alguns

estudos morfométricos aplicados aos problemas que interessam a Genética Marinha são descritos.

A forma é um fenótipo. O fenótipo é entendido como a expressão do genótipo. De maneira simples (e simplificadora) pode-se dizer (e se diz!) que o fenótipo é o genótipo mais o ambiente. Contudo, a relação entre fenótipo e genótipo é complexa e varia de característica para característica. Em um extremo, estão as características quantitativas, que possuem uma relação complexa com o genótipo. No outro extremo, o fenótipo pode ser a expressão quase que direta do genótipo. Neste caso, a distinção entre genótipo e fenótipo praticamente desaparece e pode-se dizer que, ao se observar um fenótipo, se está observando, em termos operacionais, o genótipo. Na maior parte dos estudos em Genética Marinha a variação com a qual se trabalha é aquela referente a este segundo caso, ou seja, quando se pode assumir que a variação fenotípica observada é um bom equivalente da variação gênica (como é o caso das aloenzimas, por exemplo, como foi descrito na seção 1.2 em *Controvérsias evolutivas*) ou com a própria variação gênica.

1.3.3. Variação Gênica

O conjunto de genes de uma célula haplóide de um organismo constitui o seu genoma (Em inglês: *genome* = genes, *chromosomes* = estrutura celular composta de DNA e proteínas encontrada no núcleo das células em metáfase). Toda a informação hereditária contida no genoma tem como base material o DNA (ou, em alguns vírus marinhos, no RNA). Recentemente foi observado que, além de regiões codificadoras, o DNA apresenta uma grande proporção de sequências que não são codificadoras. A definição de genoma inclui, também, estas regiões. Ou seja, o genoma pode ser entendido como todo o DNA de uma célula haplóide, seja ele codificante ou não.

No momento da divisão celular, o DNA toma uma forma semi-ordenada designada de cromossomos. Os eucariontes, como as estrelas-do-mar, possuem múltiplas fitas de DNA lineares dentro do núcleo celular, agregadas a proteínas estruturais, as histonas. Nos procariontes, como as cianobactérias, os cromossomos podem ser circulares e não possuem histonas.

As células dos organismos eucariontes podem ser classificadas de acordo com o teor cromossômico em haploides (“n” representando um conjunto de cromossomos, como no caso de algumas algas), diploides (“2n” representando dois conjuntos de cromossomos, como no maior caso dos filos com organismos marinhos) ou mais, nos casos de poliploidias (triplóides-três conjuntos de cromossomos-3n; tetraplóides-quatro conjuntos de cromossomos-4n e pentaploides-cinco conjuntos de cromossomos-5n; como no caso de alguns crustáceos que são tetraplóides e moluscos que podem apresentar diversos múltiplos de “n”).

Nas células diplóides, os cromossomos existem aos pares, ou seja, sempre existem dois cromossomos de mesmo tamanho e forma chamados, por isso, cromossomos homólogos. Cada segmento do cromossomo capaz de determinar a produção de uma proteína e, portanto, capaz de controlar uma característica morfológica ou funcional do indivíduo é conhecido por gene. Assim, a teoria cromossômica de herança definiu um cromossomo como uma sequência linear de genes (embora, atualmente, se saiba que os genes quase nunca são contínuos, com já se vai ver em *1.3.6.2. Tipos de sequências*).

O local ocupado pelo gene no cromossomo é o *locus* gênico (no plural, *loci*). Em um par de cromossomos homólogos, os genes localizados em posições correspondentes são genes alelos ou alelomorfos e atuam sempre sobre o mesmo caráter. Genes alelos (ou simplesmente alelos, para facilitar) diferentes surgem uns dos outros graças à

ocorrência de mutações gênicas (pequenas alterações na sequência de bases das suas moléculas de DNA). Dessa forma, existem diversas alternativas de alelos para um mesmo *locus*. Eventualmente, um alelo pode impedir a manifestação de outro, neste caso, ele é conhecido como alelo dominante. Aquele gene que tem o seu efeito bloqueado por um alelo dominante é chamado alelo recessivo. O fenômeno geral é conhecido como dominância e recessividade dos alelos.

Um indivíduo diplóide pode possuir dois alelos iguais em um determinado *locus* gênico e, neste caso, diz-se que ele é homozigoto para aquele *locus*. Alternativamente, quando o indivíduo apresenta alelos diferentes, diz-se que ele é heterozigoto. Para facilitar a representação do genótipo de um indivíduo se empregam, geralmente, letras para indicar os genes. Por exemplo, para um determinado *locus* com dois alelos pode se denominar um alelo como A_1 e outro como A_2 , sendo, neste caso, A_1A_1 e A_2A_2 os genótipos possíveis em homozigose e A_1A_2 o genótipo em heterozigose. O conjunto completo de alelos que podem ser encontrados no material genético dos indivíduos vivos de uma determinada população é chamado de *Pool gênico*. O *Pool gênico* para o *locus A* consiste de todos os alelos para este *locus* na população (Figura 1.11).

A variação gênica, portanto, pode ser entendida como variação aléllica. A variação aléllica pode ser descrita em vários níveis: fenotípico (característica discreta ou contínua, forma, proteínas), cromossômico ou sequências de DNA.

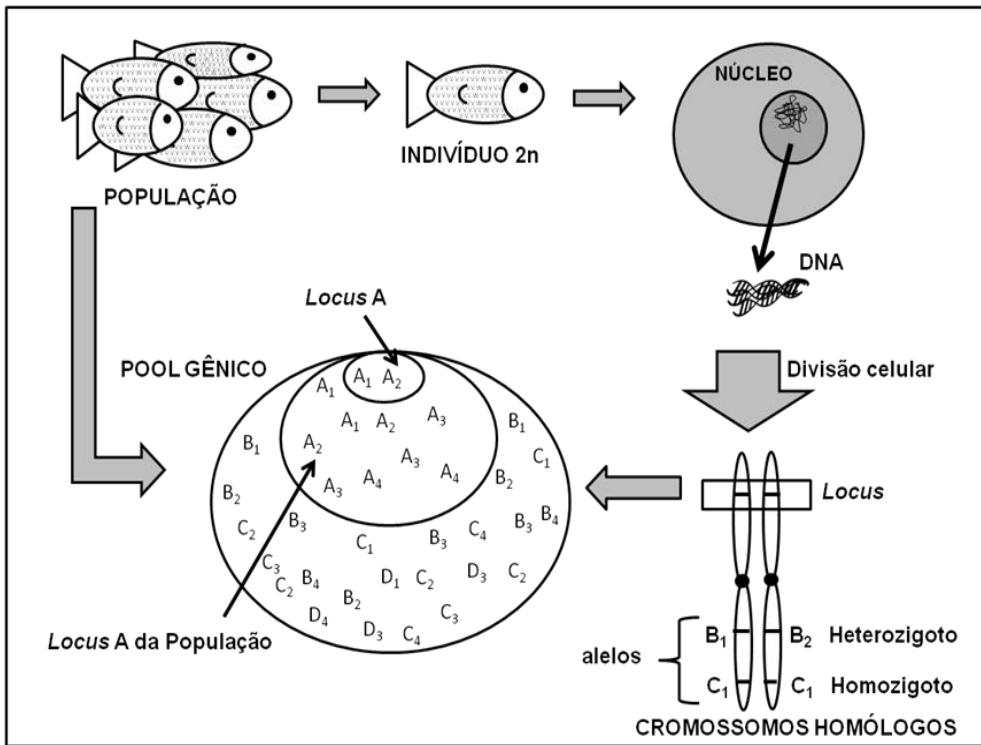


Figura 1.11. Resumo esquemático dos conceitos que servem para descrição da variação gênica em uma população (ver texto para explicação detalhada).

1.3.4. Citogenética

A citogenética é o ramo da genética relativo ao estudo dos cromossomos quanto a seus aspectos morfológicos e funcionais, sua variação e evolução (Guerra & Souza 2002; Molina & Jacobina 2013). O conjunto de cromossomos dentro de um núcleo de uma célula é chamado cariótipo.

Durante o processo evolutivo das espécies, os cariótipos se transformaram, gerando uma imensa diversidade cromossômica em número, forma (Figura 1.12), tamanho, capacidade de pareamento meiótico, conteúdo de heterocromatina etc. A análise cromossômica permite observar o genoma de um eucarionte como um todo, na forma de blocos individualizados de material genético passíveis de serem mensurados, diferenciados em sub-unidades e manipulados de diversas maneiras. A grande maioria das espécies é cariotipicamente única, diferindo de outras em relação ao número e a morfologia cromossômica e/ou em relação a características estruturais dos seus cromossomos. Caracteres cromossômicos podem ser utilizados na identificação de espécies crípticas (citotaxonomia), assim como nas relações entre espécies, fornecendo importantes subsídios para o entendimento das relações evolutivas entre espécies e populações (citossistemática). Técnicas clássicas de caracterização cromossômica (cariotipagem) incluem a coloração convencional, bandeamento C e NORs.

Os cromossomos são observados em células que se encontram em metáfase. Assim sendo, é necessário obter células em divisão para fazer um estudo cromossômico. Podem-se obter células em divisão espontânea (por exemplo, a partir das brânquias de bivalves), ou a partir de culturas cuja divisão é induzida. Os cromossomos em metáfase precisam ser corados para ser estudados. A coloração com Giemsa é a mais utilizada em técnicas citogenéticas de rotina, também chamada de coloração convencional. Este tipo de coloração é uniforme ao longo do cromossomo, não permitindo a identificação

individual (permite apenas detectar variações de número), no entanto, como deixa visível a constrição centromérica, permite identificar grupos de cromossomos de acordo com a sua morfologia e identificar alterações de largo espectro, tais como grandes variações de forma ou quebras cromossômicas.

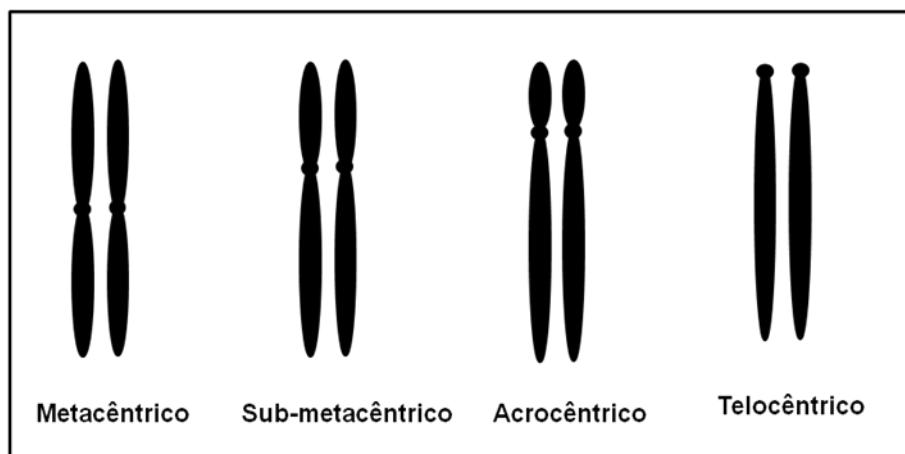


Figura 1.12. Diferentes formas de cromossomos são definidas em função da posição do centrômero. Na forma metacêntrica, o centrômero divide o cromossomo em duas partes iguais; na sub-metacêntrica, o centrômero divide o cromossomo em duas partes desiguais; na acrocêntrica, o centrômero é sub terminal e na forma telocêntrica, o centrômero é terminal.

Se o objetivo do estudo for a detecção de alterações cromossômicas, os cromossomos devem ser marcados de uma forma específica, de modo a que cada um dos pares seja identificado individualmente. Esta marcação pode ser obtida através de diferentes técnicas de bandeamento diferencial. Existem, também, métodos que marcam regiões específicas dos cromossomos e, por isso, são denominadas técnicas de bandeamento seletivo como as bandas Q, G e R.

As bandas Q foram o primeiro método de diferenciação longitudinal a ser utilizado e baseia-se na coloração dos cromossomos pela mostarda de quinacrina e observação da fluorescência emitida usando luz ultravioleta. Os cromossomos são corados sem qualquer tipo de pré-tratamento, preservando, assim, a sua morfologia. Esta técnica apresenta, contudo, uma desvantagem: a fluorescência desvanece

rapidamente. Assim sendo, para a análise citogenética de rotina este método foi substituído por técnicas de bandeamento não fluorescente.

As bandas G podem ser obtidas fazendo um pré-tratamento dos cromossomos com uma solução salina a 60°C ou com enzimas proteolíticas, seguido de coloração com Giemsa. Os cromossomos apresentam, então, um padrão de bandas claras e escuras, específico para cada cromossomo. As bandas R (*reverse*, do inglês) são o reverso das bandas G. A técnica, neste caso, envolve um pré-tratamento alcalino a temperatura elevada (80 a 90°C), seguido de coloração com Giemsa.

As bandas mostradas pelas técnicas de bandeamento refletem a organização funcional do genoma que regula a replicação do DNA, a reparação, a transcrição e a recombinação genética. As bases moleculares das técnicas de bandas envolvem a composição de bases do DNA, as proteínas associadas e a organização funcional do genoma. De um modo geral, as bandas G escuras (R claras) são ricas em bases AT, replicam tarde e são pobres em genes. As bandas G claras (R escuras) são ricas em bases CG, replicam precocemente e são ricas em genes.

Os cromossomos em prometáfase podem, também, ser bandeados. Neste caso, a técnica utilizada é o bandeamento C. Este tipo de marcação permite aumentar o número de bandas visíveis e localizar, com maior rigor, alterações cromossômicas sutis, difíceis de serem detectadas em cromossomos em metáfase. A heterocromatina constitutiva que está localizada nas regiões centroméricas dos cromossomos é marcada seletivamente, por isso esta técnica é aplicada ao estudo dos polimorfismos que envolvem uma região denominada justa-centromérica. As bandas C podem ser obtidas através de vários métodos, mas o mais comum consiste num breve tratamento com ácido, seguido de tratamento alcalino e, posteriormente, coloração com Giemsa.

As bandas NOR correspondem à marcação das regiões dos organizadores nucleolares (*nucleolar organizing regions*, em inglês) que se encontram, geralmente, nos braços curtos dos cromossomos acrocêntricos. O método mais utilizado é o da precipitação de nitrato de prata sobre aquelas regiões. A precipitação não se dá sobre o DNA ribossômico, mas resulta da redução de uma (ou várias) proteínas ácidas argirófilas que se encontram associadas àquelas regiões, onde se localizam os genes que codificam as subunidades 18S e 28S do rRNA (ver adiante em *1.3.6.1. DNA alvo*). O precipitado localiza-se, por isso, no exterior dos cromossomos e só aparece se os referidos genes estiverem ativos.

A citogenética contribui para resolução de muitos problemas de Genética Marinha como em casos em que a taxonomia é controversa. A classificação das espécies do gênero *Carangoides*, conhecidas popularmente como Xerelete, Carapau e Xaréu, puderam ter sua classificação resolvida através da cariotipagem das regiões 18S e 5S do rRNA. Na seção **1.6. Científico**, alguns estudos com citogenética são descritos.

1.3.5. Eletroforese de Aloenzimas

Na seção 1.2 deste capítulo, foi abordado, em dois momentos, o método de eletroforese de aloenzimas. Primeiro, na história da teoria evolutiva, a eletroforese de aloenzimas foi apresentada como o método que deu fim a controvérsia entre a Escola Clássica e a do Balanço (*1.2.1.3. Controvérsias evolutivas*). Depois, na história da própria Genética Marinha (**1.2.2. O Começo**), fez-se referência a este método como sendo importante para o nascimento e crescimento da Genética Marinha. Agora, chegou o momento de descrever este método em mais detalhes.

1.3.5.1. A técnica

O desenvolvimento da técnica de eletroforese de proteínas data da década de 1950 quando Smithies publicou um trabalho (*Zone electrophoresis in starch gels: group variations in the serum proteins of normal human adults*) que permitia uma boa resolução de proteínas separadas por corrente elétrica em gel de amido (Smithies 1955) e Hunter e Market desenvolveram métodos de coloração histoquímica baseados na especificidade enzimática (*Histochemical demonstration of enzymes separated by zone electrophoresis in starch gels*) (Hunter & Market 1957).

A eletroforese pode ser definida, de maneira geral, como a migração de partículas sob ação de corrente elétrica. No caso da eletroforese de proteínas, a técnica se baseia, primeiramente, nas características físico-químicas das proteínas. Ou seja, é sabido que dos vinte aminoácidos, cinco possuem carga elétrica, três dos quais (lisina, arginina e histidina) positiva e, os demais (ácido glutâmico e ácido aspártico), negativa. Tal fato faz com que proteínas com diferentes sequências de aminoácidos possam apresentar cargas diferentes, movimentando-se com diferentes velocidades, sob a ação de um campo elétrico. Em outras palavras: diferentes proteínas podem manifestar diferentes mobilidades eletroforéticas (Avise 1974; Avise 1994; Murphy et al. 1996; Solferini & Selivon 2012).

As proteínas possuem diferentes funções nos organismos, podendo ser estruturais (como as proteínas que participam da arquitetura celular, conferindo formas, suporte e resistência) ou enzimas (proteínas que são catalisadores biológicos com alta especificidade). Devido ao fato de serem catalisadoras, as enzimas podem ser visualizadas nos géis de eletroforese com a utilização de coloração histoquímica.

Para se realizar uma eletroforese de enzimas, algumas etapas básicas são necessárias. São elas: preparação do suporte, preparação dos extratos, corrida eletroforética e coloração histoquímica.

O suporte para a eletroforese pode ser de amido, poliacrilamida, agarose ou acetato de celulose, sendo que amido e poliacrilamida permitem uma melhor separação das enzimas. Os géis podem ser preparados com diferentes tampões, que são escolhidos de acordo com a enzima e o organismo utilizado. Cada diferente tampão utilizado para fazer os géis possui um tampão que será utilizado nos eletrodos. Os dois tampões juntos constituem o sistema de tampões. Os sistemas de tampões podem ser contínuos, quando o tampão utilizado para preparar o gel é o mesmo utilizado nos eletrodos ou descontínuos, quando o tampão utilizado para preparar o gel é diferente daquele utilizado nos eletrodos.

A extração das enzimas é feita a partir da homogeneização de indivíduos inteiros (caso de organismos pequenos com sementes de mexilhões, por exemplo), órgãos ou tecidos (caso de organismos maiores, como peixes). A homogeneização precisa ser realizada em ambiente refrigerado com o auxilio de um bastão de vidro em uma placa mantida sobre gelo para evitar a proteólise e desnaturação das enzimas. O extrato de cada indivíduo pode ser aplicado diretamente no gel através de pedaços de papel de filtro de aproximadamente 2 x 15mm mergulhados no homogeneizado e logo em seguida colocados nos géis.

O gel com os extratos das enzimas precisa ser submetido a uma corrente elétrica, processo chamado de corrida eletroforética. As condições, tanto de voltagem aplicadas ao gel como de duração de tempo de corrida variam de acordo com o sistema de tampão utilizado. As corridas podem durar de 3 a 24 horas a uma potência média de 3,2 W. As

corridas eletroforéticas devem ser realizadas em um ambiente com baixa temperatura, bem como todas as outras etapas do processo.

Após a corrida eletroforética, os géis são colocados em placas onde a coloração histoquímica (ou soluções reveladoras) apropriada para cada enzima é adicionada. As soluções reveladoras consistem de um tampão com pH próximo do que é considerado ideal para a enzima, onde são dissolvidos o substrato da enzima, seus cofatores e substâncias capazes de detectar a reação ou formação do produto que precipitam, oxidam ou fluorescem em consequência da reação principal.

Os padrões de bandas resultantes da revelação do gel (zimograma) devem ser interpretados seguindo a estrutura quaternária das proteínas mono, di, e tretraméricas. Os alelos são classificados de acordo com sua mobilidade anódica. Contudo, a interpretação de um zimograma como a representação de fatores hereditários mendelianos depende de um salto qualitativo, que é a transformação desta técnica em um método de amostragem de variação gênica.

1.3.5.2. O método

Em 1966, Dick Lewontin e Jack Hubby definiram as condições necessárias para que uma técnica fosse capaz de amostrar o genoma de um organismo possibilitando, assim, uma medida da variação gênica de uma população natural. Segundo eles, esta técnica deveria detectar diferenças fenotípicas decorrentes da substituição alélica em um único *locus* (1); distinguir substituições alélicas de um *locus* em relação a outros *loci* (2); discriminar uma quantidade substancial das substituições alélicas (3) e, finalmente, representar uma amostra imparcial do genoma em relação aos efeitos fisiológicos e grau de variação (4). As condições 1 e 2 garantiriam que os fenótipos observados apresentariam uma herança mendeliana simples e sem variação ambiental importante.

As condições 3 e 4 dariam a segurança de que, a despeito de serem obtidos a partir de uma amostra restrita, os resultados poderiam ser usados para se fazer afirmações sobre a variação no genoma como um todo.

Lewontin e Hubby indicaram que, se o dogma central da biologia molecular (Figura 1.13) fosse aceito como verdadeiro, seria possível interpretar a variação de isoenzimas (enzimas com mobilidade diferente, mas mesma função) observada nos zimogramas como equivalente da variação gênica. Assim, a técnica de separação de enzimas, se tornava um método para amostragem da variação gênica em populações naturais. O pressuposto era, já naquele momento, falso. Muita variação genômica não é detectável por eletroforese simplesmente porque mudanças no DNA que não mudem a codificação do aminoácido (substituições em terceiras bases dos códons, por exemplo), mudanças de aminoácidos que não mudem a carga da proteína (mudança de um aminoácido para outro com a mesma carga- Lisina para arginina, por exemplo) e mudanças em regiões não codificadoras do genoma (em introns, por exemplo) não são detectadas pelo método. Contudo, todas estas limitações faziam da eletroforese de aloenzimas, apenas, um método conservador para responder a pergunta sobre a variação gênica presente nas populações naturais. Ser conservador, em ciência, não é uma coisa de todo mal. Isto garante que toda afirmação que for feita sobre a existência de variação é muito segura. Por outro lado, toda afirmação sobre a falta dela esta sujeita as limitações do método.

Voltando aos padrões de banda resultantes da revelação do gel, ele, agora, pode ser interpretado mendelianamente. Cada banda representa um alelo, que é classificado de acordo com sua mobilidade anódica. O alelo mais rápido de cada *locus* é designado “A”, o segundo “B”, e assim por diante. A interpretação é simples: indivíduos que apresentem apenas uma banda no gel são interpretados como homozigotos (AA, BB,

CC etc.). Indivíduos que apresentem as duas formas da enzima, com mobilidades eletroforéticas diferentes (isoenzimas), são interpretados como heterozigotos (AB, AC, BC). A figura 1.14 representa as etapas do método de eletroforese de izoenzimas que, se interpretadas mendelianamente, passam a ser chamadas aloenzimas (isoenzimas com padrão mendeliano de herança).

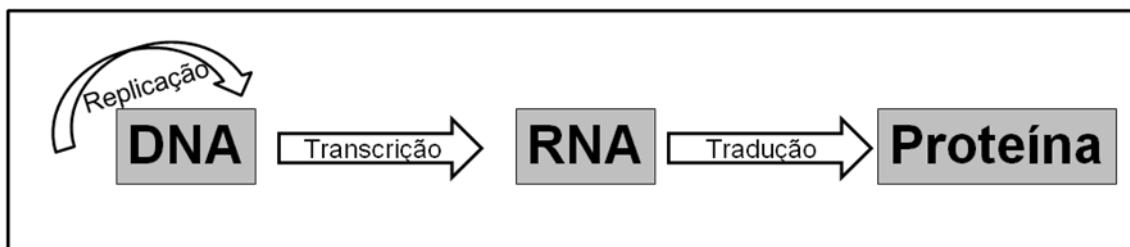


Figura 1.13. Dogma Central da Biologia Molecular descrito pela primeira vez em 1958, por Francis Crick, um dos responsáveis pela proposição do famoso modelo de dupla hélice para o DNA. Em 1970 Crick voltou ao já popular Dogma em um artigo para *Nature*.

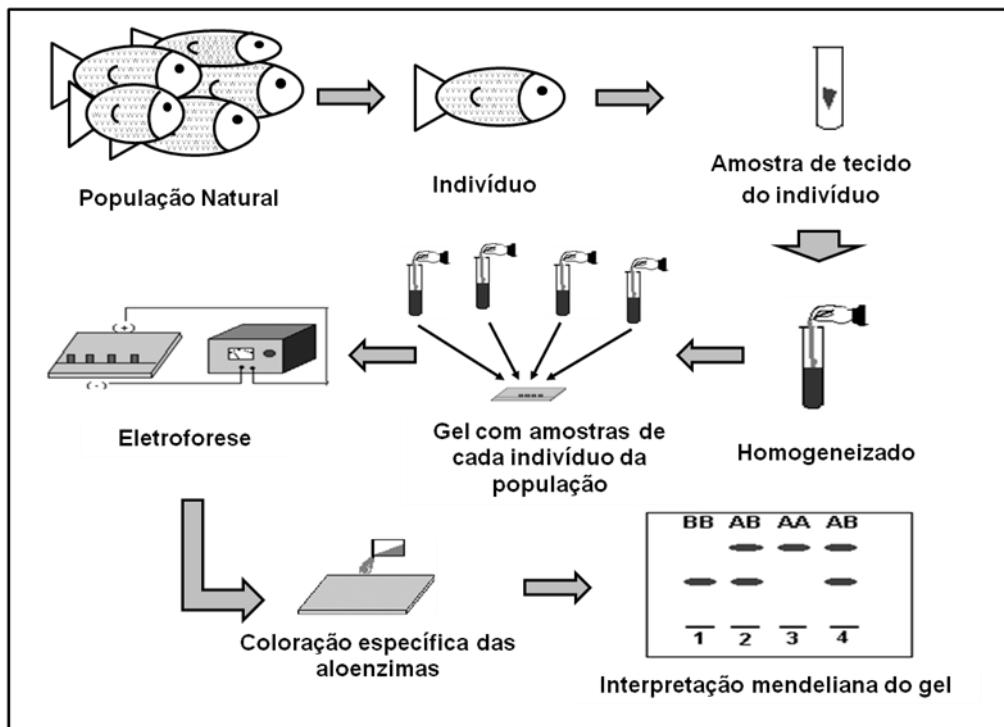


Figura 1.14. Resumo do método de eletroforese de aloenzimas tendo, ao final, a interpretação mendeliana do padrão de bandas.

O uso da eletroforese de aloenzimas como método de amostragem da variação gênica de populações naturais incorre, no entanto, na aceitação de algumas limitações. Talvez, a maior delas seja o fato de que a variação amostrada por eletroforese seja, na realidade, variação fenotípica e não genotípica. Duas consequências advêm daí. Primeiro que conceitos como heterozigosidade e homozigosidade ficam relativizados, já que se trabalha com produtos dos genes. Segundo que os chamados alelos (melhor seria a designação técnica de alelomorfos) são afetados não só por possíveis alterações pós-transcricionais, como também pela ação do ambiente e, ainda, por condições de estocagem do material a ser analisado (tempo, refrigeração etc.).

Para além das suas limitações, a utilização do método de eletroforese de aloenzimas oferece muitas vantagens como, por exemplo, fácil preparação de extratos a ser utilizados. Acima de tudo, a interpretação mendeliana das frequências obtidas em seus padrões de bandas é uma das maiores vantagens da eletroforese. Desse modo, este método é, reconhecidamente, um dos métodos mais eficazes na detecção e amostragem de variação gênica em populações e vem sendo usado com sucesso em estudos de Genética Marinha desde a década de 1970. A partir da década de 2000, contudo, a utilização deste método tem sofrido um declínio que tem suas possíveis razões discutidas mais a frente (ver *1.3.8. E Agora José?*).

1.3.6. Técnicas de DNA

A Genética Marinha só pôde ter acesso à variação gênica presente nas sequências de DNA a partir da década de 1970. Contudo, como marcaria David Skibinski em um artigo de 1994 (*The potential of DNA techniques in the population and evolutionary genetics of aquatic invertebrates*), o impacto do desenvolvimento das

técnicas de DNA só começaria a ser sentido na área a partir da década de 1990 (Skibinski 1994).

A história dos estudos com DNA começa na década de 1950, como é bem sabido. A partir dos resultados de difração de raios X, Rosalind Franklin e Maurice Wilkins, em 1952, deduziram que as moléculas de DNA eram helicoidais. O modelo tridimensional para a estrutura da molécula de DNA foi proposto, em 1953, por James Watson e Francis Crick (Watson & Crick 1953). A enzima polimerase do DNA foi descrita em 1955, por Arthur Kornberg. No final da década de 1950, foi demonstrada a natureza tríplice do código genético – cada aminoácido de uma proteína é especificado por uma sequência de três pares de nucleotídeos no DNA. Logo em seguida, decifrou-se o código genético e demonstrou-se a correspondência entre as trincas de bases nitrogenadas do mRNA (códons) e os aminoácidos nas proteínas. O código é inequívoco, cada códon corresponde a um único aminoácido, e degenerado, pois quase todos os aminoácidos têm mais de um códon que o codifica.

Para ser trabalhado, o DNA precisa ser extraído das células. Isto se faz, geralmente, com uma etapa de tratamento dos tecidos com enzimas proteolíticas e detergentes aniônicos e, em sequência, um tratamento com fenol e clorofórmio e precipitação com álcool. Na década de 1970, a tecnologia do DNA recombinante trouxe a possibilidade de extrair, cortar e unir DNA exógeno em vetores, como fagos e plasmídeos, a partir do isolamento das enzimas de restrição.

As enzimas de restrição (ou endonucleases de restrição) foram isoladas, em 1968, por dois grupos independentemente (*Host specificity of DNA produced by Escherichia coli*, Linn & Arber 1968 e *DNA restriction enzyme from E. coli*, Meselson & Yuan 1968). Estas enzimas são capazes de digerir sequências específicas de quatro, cinco ou seis pares de bases em DNA dupla fita e, desta forma, abriram o caminho para

formação de moléculas híbridas (sequências de DNA das espécies que se queria estudar unidas a sequências de bactérias ou vírus que, por reprodução, formariam várias cópias da sequência desejada) que passaram a ser de extrema utilidade no estudo da variação.

O desenvolvimento mais importante da biologia molecular para a Genética Marinha, contudo, foi a PCR (*Polimerase Chain Reaction*) desenvolvida, na década de 1980, por Saiki et al. (1988) (*Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase*). Naquele momento, tornava-se possível, para um grande número de laboratórios, trabalhar com sequências de DNA, tanto nucleares quanto organelares.

1.3.6.1. DNA alvo

Nos eucariontes, as moléculas de DNA podem ser encontradas tanto no núcleo das células quanto nas suas organelas (mitocôndrias e cloroplastos). O DNA nuclear (nDNA) é composto, em geral por moléculas bastante longas, podendo atingir vários centímetros de comprimento. A maior parte deste DNA, no entanto, nunca é transcrita em RNA e não contém informação para a produção de proteínas. Calcula-se que apenas cerca de 2% do DNA de eucariontes corresponde aos genes. Os 98% restantes são sequências de nucleotídeos que não produzem RNA e cuja função ainda não é completamente conhecida, constituindo o DNA não-codificante (Bayne 1996; Burton 1996; Nahum 2012).

Embora a maior proporção do DNA dos organismos eucariontes seja encontrada nos cromossomos nucleares, existem duas organelas celulares – as mitocôndrias e os cloroplastos – que apresentam tipos particulares de cromossomos. Estes cromossomos contêm genes que codificam funções específicas dessas organelas. O cromossomo

mitocondrial (mtDNA) e o cromossomo do cloroplasto (cpDNA) apresentam número de genes reduzido em relação aos nucleares.

Os cromossomos de cloroplastos e mitocôndrias possuem algumas características que os diferem dos cromossomos nucleares. Em geral, são circulares, embora existam evidências de que podem adquirir formas lineares. Possuem um grau de condensação menor que os nucleares e se encontram presentes nas células em centenas ou milhares de cópias. O cpDNA é dezenas de vezes maior que o mtDNA e, estes, de uma a seis ordens de grandeza menores que o nDNA.

Quanto ao modo de herança, tanto o mtDNA quanto o cpDNA apresentam herança citoplasmática, sendo que o mtDNA é herdado, geralmente, via materna, já o cpDNA via paterna e, também, biparental. Para ambas as moléculas não existem evidências de recombinação. Já o nDNA apresenta herança mendeliana e a recombinação é característica no modo de herança biparental desta molécula. Em relação às taxas evolutivas, o mtDNA evolui em torno de dez vezes mais rápido que o nDNA, devido, provavelmente, ao fato da replicação em mitocôndrias não apresentar um sistema próprio de reparo e sua função respiratória gerar muito radicais livres que são danosos a molécula de DNA. O cpDNA, por sua vez, apresenta as menores taxas evolutivas comparadas ao mtDNA e, mesmo, ao nDNA. Neste sentido, o cpDNA não é, geralmente, a melhor escolha para amostragem de variação (Silva & Russo 2000).

Tanto o DNA nuclear quanto o organelar possuem genes que codificam os RNAs ribossômicos, que são componentes essenciais na fisiologia celular. Estes componentes interagem de modo específico com proteínas ribossômicas para formar as subunidades dos ribossomos que atuam na síntese de proteínas. Os rRNAs são o principal produto da transcrição, constituindo, geralmente, de 80 a 90% da massa de RNA total dos procariontes e eucariontes (Gorab 2012).

As sequências que codificam para rRNA são reiteradas, ocorrendo em número variável nos diversos organismos estudados. Nos eucariontes, essas cópias são organizadas *in tandem* (isto é, em sequência) e agregados a uma ou mais regiões cromossômicas, as chamadas regiões organizadoras de nucléolos. De forma geral, cada unidade de repetição do rDNA eucarionte possui uma organização conservada, consistindo de um espaçador externo (ETS, do inglês *external transcribed spacer*), transcrito em uma sequência que contém a extremidade 5' da molécula precursora do RNA; uma região que codifica para rRNA 18S; um espaçador interno que é transcrito (ITS, do inglês *internal transcribed spacer*) e que contém a região que codifica para rRNA 5,8S; uma região que codifica para rRNA 25-28S e um espaçador externo (NTS, do inglês *non transcribed spacer*) que não é transcrito.

Como pode ser percebido, o genoma possui grande variação quanto ao seu tamanho, organização e composição. Essa variação reflete a dinâmica da evolução dos genomas, sendo relevante em estudos da biodiversidade molecular e fundamentais para os estudos com Genética Marinha. Na figura 1.15 estão esquematizadas a estrutura de três das formas de DNA aqui descritas (nDNA, mtDNA, rDNA) e que servem como alvos para amostragem de variação gênica para os estudos de Genética Marinha.

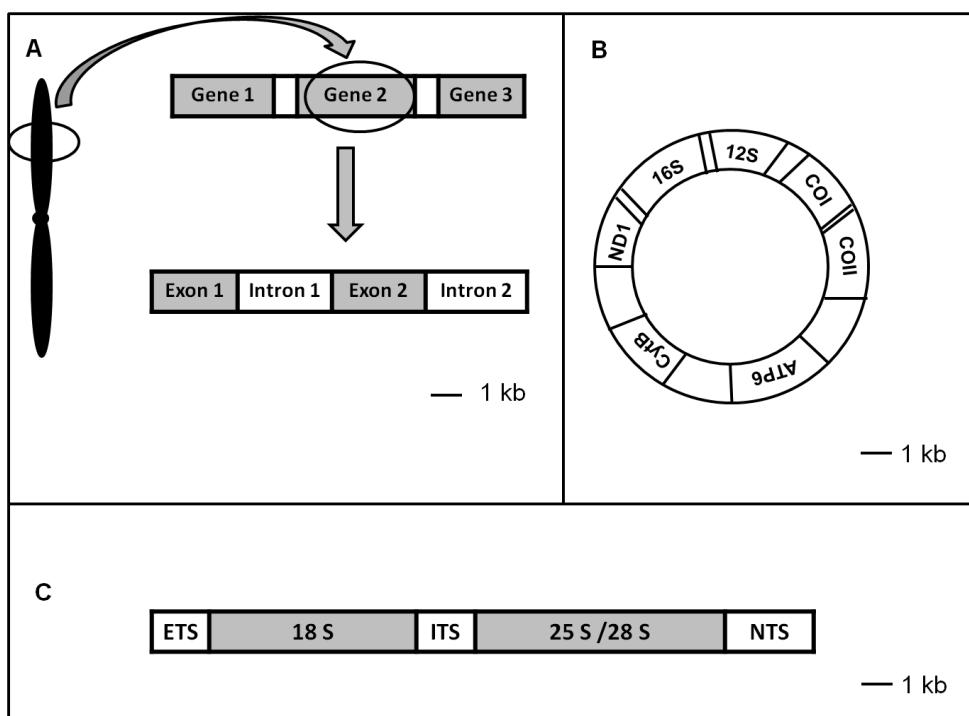


Figura 1.15. Representação esquemática das três formas de DNA. A- genoma nuclear, nDNA, B- genoma mitocondrial, mtDNA (com os genes: COI, COII, ATP6, CytB, ND1, 16S e 12S) e C- unidade de repetição do rDNA de eucariontes (ETS: espaçador externo, 18S: região que codifica para rRNA 18S, ITS: espaçador interno, 25S/28S: região que codifica para rRNA 25/28S, NTS: espaçador externo não codificante ou espaçador intergênico). kb – kilobase. 1 kb é equivalente a 1000 pb (pares de bases).

1.3.6.2. Tipos de sequências

As regiões codificantes do genoma é que constituem os genes. Contudo, os genes não são dispostos de maneira contínua nas moléculas de DNA, eles se encontram interrompidos por sequências não codificantes. As sequências codificantes formam os chamados exons que são intercalados por sequências não codificantes, os chamados introns (ou regiões intergênicas). A constituição de um gene se dá após o processo de transcrição por um processamento, o chamado *splicing* (em inglês), que retira os introns transcritos e une os exons que estavam por eles separados.

Os exons são mais conservados (ou seja, evoluem de maneira mais lenta), o que significa dizer que a quantidade de mudanças que eles acumularam ao longo do tempo é

pequena. Assim, exons não apresentam muita variação dentro de uma mesma espécie. Os introns, por outro lado, como são não codificantes, evoluem mais rápido e, portanto, apresentam grande quantidade de variação entre indivíduos de uma mesma espécie. A razão pela qual os exons são conservados e os introns não são é, provavelmente, porque exons são os componentes dos genes e, portanto, se sofrerem mudanças (as mutações, que são sempre ao acaso) podem perder sua função e causarem danos (ou até a morte) aos indivíduos. Dito de outra forma, os exons estão sob seleção natural e isto os impede de mudar, a não ser que a mudança seja benéfica ou neutra. Os introns, como não codificam nada, estão livres para ir mudando ao longo do tempo sem que nada aconteça aos indivíduos, ou seja, estão livres de seleção natural. Se o objetivo é amostrar a variação em nível populacional, os introns, certamente, constituem a melhor escolha.

Além dos exons e introns, o nDNA contém, também, regiões de cópia única e regiões duplicadas (nDNA repetitivo). As regiões de cópia única geralmente codificam um determinado produto do gene. O nDNA repetitivo é composto de sequências que se repetem em diferentes graus, podendo ser constituídos por segmentos de codificação, tais como o rDNA que codifica o RNA ribossomal (rRNA) ou não-codificante, *in tandem*. As unidades repetitivas incluem algumas das regiões mais variáveis identificadas no genoma de eucariontes.

As unidades repetitivas estão ligadas entre si como repetições em sequências e podem ser encontradas em um *locus* ou em muitas regiões espalhadas por todo o genoma. Essas repetições de número variável em sequência (VNTRs, do inglês *Variable Number of Tandem Repeats*) são chamadas de DNA satélite (Jeffreys et al. 1985a, 1985b). A variação gênica presente nestes segmentos de DNA podem ser do tipo microsatélites (repetições constituídas de 2 a 6 pares de bases) e minisatélites (até 10, 20

pares de base de repetição). As sequências de rDNA, por outro lado, são mais conservadas.

O genoma mitocondrial também inclui uma pequena região de DNA não-codificante conhecida como laço de deslocamento (D- Loop, do inglês *Displacement Loop*), que serve de origem de replicação para o genoma mitocondrial. Na maioria dos animais, o D-Loop é muito mais variável do que o resto do genoma mitocondrial.

Os diferentes níveis de variação das sequências de DNA podem ser usados para estudar diferentes problemas em Genética Marinha como identificação de paternidade, diferenciação entre populações e relações de ancestralidade. Por serem usados como rastreadores ou marcas do processo evolutivo das populações (ou seja, como elas vem mudando ao longo do tempo e do espaço) estas sequências, e também as aloenzimas, constituem os famosos marcadores moleculares (moléculas de proteínas e DNA) em contraposição aos marcadores cromossômicos e fenotípicos (quantitativos, qualitativos e de forma).

Dados os tipos de genoma e sequências, a questão agora é como acessar a variação. Diferente da eletroforese de aloenzimas, a variação em nível de DNA depende de uma série de técnicas que possibilitem, depois da extração do DNA das células, a inferência da variação gênica e a sua visualização. A visualização depende, geralmente, também, de uma eletroforese e métodos de coloração, como na eletroforese de aloenzimas. As técnicas de inferência, contudo, são diversas e constituem uma verdadeira “sopa de letrinhas”. Para começar o RFLP.

1.3.6.3. RFLP

Mudanças de nucleotídeos em sequências do DNA (as chamadas mutações pontuais) podem ser inferidas diretamente em genomas pequenos, como o mtDNA, pelo

uso de enzimas de restrição. Estas enzimas cortam as sequências toda vez que encontram segmentos específicos de DNA. Por exemplo, a enzima *MboI* reconhece e corta a sequência GATC e a enzima *BamHI* reconhece e corta a sequência GGATCC.

O resultado da atividade destas enzimas é a produção de segmentos de DNA que vão variar de tamanho em função da presença ou ausência destes segmentos específicos, os chamados sítios de restrição. Assim, feita uma eletroforese com o mtDNA após a ação destas enzimas, varias bandas serão evidenciadas após a coloração. Este padrão de bandas que varia entre os indivíduos de uma população, na dependência das sequências específicas de cada um, forma o chamado “Polimorfismo de Tamanho dos Fragmentos de Restrição” ou RFLP (do inglês, *Restriction Fragment Length Polymorphism*) (Clark 1987; Clark 1990; Arias & Infante-Malachias 2012).

Diferentes enzimas de restrição podem ser utilizadas e cada uma reconhece uma sequência única de quatro, cinco, seis ou oito bases. A escolha de uma enzima pode ser efetuada considerando o tamanho e o número de fragmentos de DNA desejado, seguido por tentativa e erro para encontrar a melhor enzima para um sistema particular. O uso mais extenso de enzimas de restrição em Genética Marinha tem sido observado em pesquisas de diversidade de alelos e diferenciação em mtDNA de populações animais. Isto porque a digestão enzimática da pequena molécula de mtDNA resulta em um pequeno número de fragmentos. Os padrões de bandas podem ser identificados como referentes a diferentes alelos ou haplótipos (designação derivada de fato de que o mtDNA é uma molécula única, portanto, apresenta apenas o número haplóide). A figura 1.16 mostra o racional que está na base da interpretação da técnica de RFLP e seus padrões de bandas.

Moléculas maiores, tais como o nDNA, não pode ser analisadas diretamente por enzimas de restrição porque existem muitos sítios de clivagem enzimáticas para ser

interpretável, os padrões a serem observados de nDNA clivado aparece como uma mancha na eletroforese. Desta forma, fragmentos de interesse particular devem ser amplificados através da técnica de PCR.

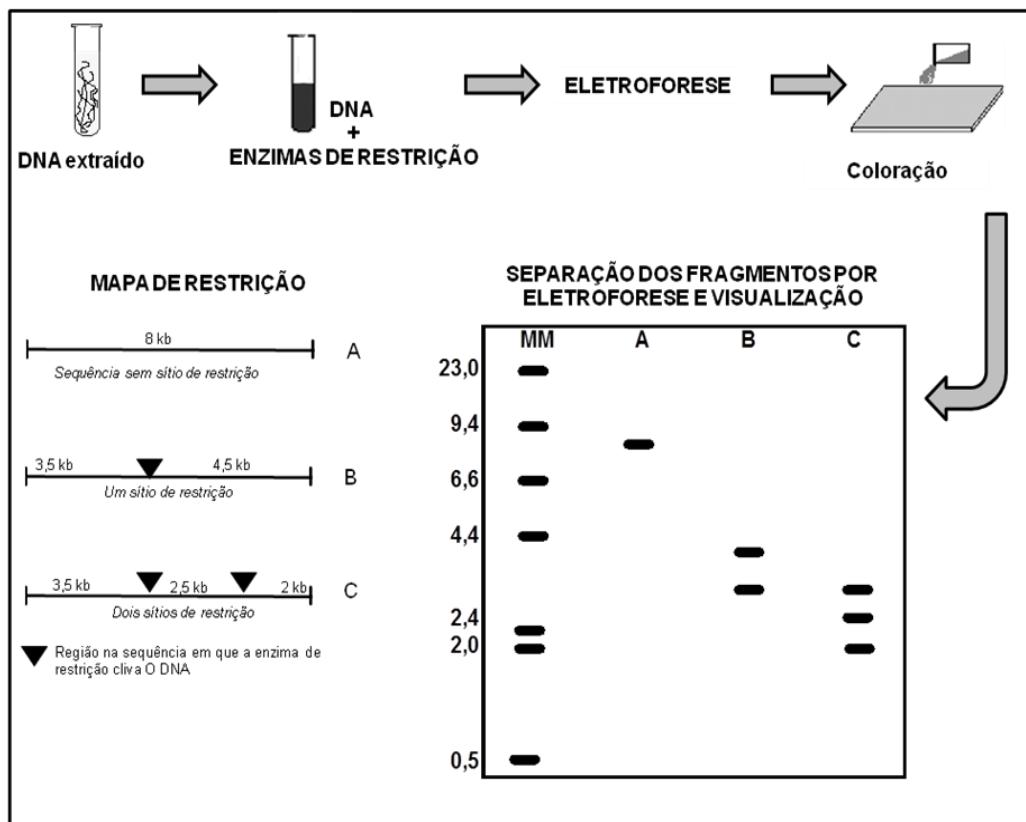


Figura 1.16. Representação esquemática da técnica de RFLP. Seguindo as setas: DNA extraído, adição das enzimas de restrição ao DNA extraído, eletroforese, coloração e padrão de bandas revelando o polimorfismo dos fragmentos de restrição interpretados no mapa de restrição ao lado. MM é o marcador molecular de tamanho. O alelo A é a sequência de DNA de 8 kb que não sofreu restrição, portanto, no gel é possível observar apenas uma banda no tamanho de 8 kb, como evidenciado pelo MM. O alelo B apresentou apenas um sítio de restrição, portanto, a sequência de 8 kb foi cortada em dois fragmentos (3,5 kb + 4,5 kb) como mostrado no gel, podendo seu tamanho ser conferido em relação ao MM. O alelo C, com dois sítios de restrição, originou os três fragmentos mostrados no gel (3,5 kb + 2,5 kb + 2,0 kb). Lembrando que kb = kilobase que equivale a 1000 pb (pb = pares de bases).

1.3.6.4. PCR

O isolamento da polimerase de DNA termicamente estável a partir da bactéria *Thermus aquaticus* (*Taq* polimerase) levou a um meio eficiente de amplificar

fragmentos curtos de DNA. A PCR ocorre em uma máquina chamada termociclador que repete vários ciclos de três temperaturas: a primeira para desnaturar o DNA e parar todas as reações enzimáticas; a segunda para que os iniciadores possam se unir à fita molde de DNA e, a última, para que o processo enzimático de replicação do DNA seja levado a cabo. O ambiente químico desta reação contém, iniciadores sintéticos de oligonucleotídeos (geralmente 20-30 bases de comprimento, os chamados *Primers*, em inglês) que são complementares para cada uma das regiões de flanqueamento. Além dos iniciadores e da fita molde de DNA, a enzima que polimeriza a reação (*Taq* DNA polimerase), o material para construir as novas fitas de DNA (dNTPs, que são os quatro nucleotídeos necessários para polimerização da nova fita de DNA) e um co-fator (ou mais) que auxilia a reação (geralmente MgCl₂, cloreto de magnésio). Tudo isto dissolvido em água destilada, filtrada e estéril. A cada ciclo completo de reação, o processo se repete, o que proporciona a replicação exponencial do DNA. Ao final do processo se obtém um grande número de cópias do segmento de DNA flanqueados pelos iniciadores (Matioli 2001; Matioli & Fernandes 2012; Matioli & Passos-Bueno 2012). Na figura 1.17 as etapas que constituem a técnica de PCR para o estudo de Genética Marinha é resumida.

Para a maioria das utilizações da PCR é preciso determinar as sequências das regiões flanqueando um determinado *locus*. Isto pode implicar um esforço considerável quando se trabalha com uma espécie para qual não existem informações sobre as suas sequências de DNA. O conhecimento das regiões flanqueadoras, por outro lado, permite a amplificação de sequências de todos os tipos e genomas: mtDNA, cpDNA, rDNA, minisatélites e microsatélites.

A amplificação de introns constitui a técnica conhecida com EPIC (do inglês, *Exon-Primed Intron-Crossing*). Neste caso, utilizam-se *primers* que anelam nas regiões

exônicas, que são conservadas e flanqueiam os introns de interesse. Como os introns não tem papel direto na tradução gênica, eles têm tendência de acumular mutações, desta forma sendo grandes repositórios de variação gênica (Côrte-Real et al. 1994a; 1994b).

A parte a amplificação destes tipos de genoma e sequências, a primeira aplicação da técnica de PCR para detecção de polimorfismos foi a sua combinação com a técnica de RFLP (PCR-RFLP). O sucesso na obtenção de resultados com está técnica depende, em primeiro lugar, da amplificação da região pretendida e, depois, da existência de sítios de restrição na região amplificada. A técnica de PCR-RFLP tem uma limitação com relação à amplificação de fragmentos grandes e, portanto, quanto menor o fragmento menor a chance de se encontrar um número expressivo de sítios de corte para as enzimas de restrição.

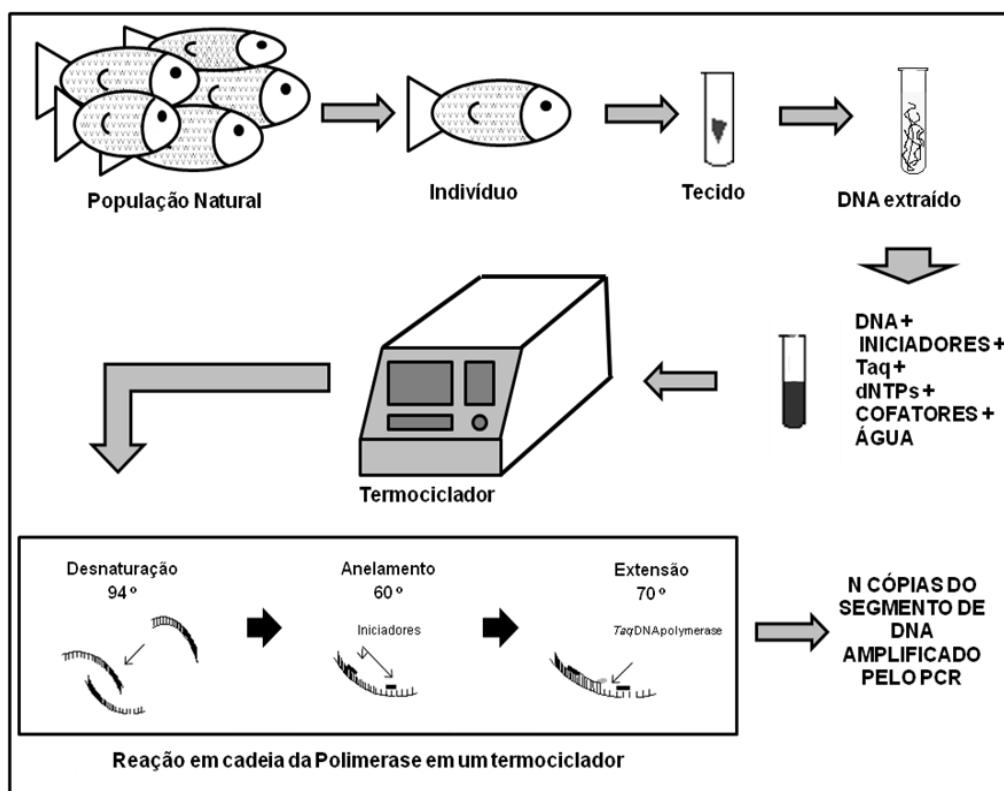


Figura 1.17. Representação esquemática das etapas que constituem a técnica de PCR.

Existem técnicas que permitem o estudo de mutações pontuais em fragmentos de DNA. Algumas destas técnicas baseiam-se nas diferenças de mobilidade eletroforética de fitas simples de DNA enoveladas em conformação tridimensional. As diferentes conformações tridimensionais das fitas simples de DNA se devem a sua sequência de nucleotídeos. Esta técnica é conhecida como SSCP (do inglês, *Single Strand Conformational Polymorphism*). Para se obter as fitas simples, o DNA é aquecido a temperaturas elevadas (maiores que 90 °C) e impedidas de re-naturação pela adição de algumas substâncias como o formaldeído. Um aspecto interessante nesta técnica é o seu padrão de bandas. Diferentemente da eletroforese de aloenzimas, homozigotos apresentam duas bandas e, heterozigotos, quatro bandas no gel. Isto se deve ao fato de que sequências de nucleotídeos iguais nas duas duplas fitas desnaturadas, apresentarão duas conformações espaciais: uma para as duas fita 5'-3' e, outra, para a duas fitas 3'-5'. Heterozigotos apresentarão quatro conformações espaciais, uma para cada uma das quatro fitas simples.

Alternativamente ao conhecimento das sequências, é possível se utilizar técnica de PCR para amplificar segmentos aleatórios de DNA ao longo de todo o genoma dos organismos. Nesta estratégia, iniciadores pequenos (por volta dos 10 pares de base) e de sequência aleatória são utilizados para a PCR com temperaturas que permitam o seu pareamento em diversas regiões do genoma, amplificando desta forma um grande número de segmentos de DNA dos mais diversos tamanhos e origens. Esta técnica é conhecida como RAPDs (do inglês *Randomly Amplified Polymorphic DNA's*). Um dos problemas dessa técnica diz respeito à falta de reproduzibilidade dos resultados, pois as condições de hibridação para detecção de polimorfismo têm que ser discriminantes dentro de um espectro muito curto das condições experimentais. O aumento no tamanho

dos *primers* asseguraria uma melhor reproduzibilidade, mas isto acarreta na diminuição da probabilidade de produção de bandas.

Uma técnica chamada de SPAR (do inglês, *Single Primer Amplification Reaction*) ou ISSR (do inglês, *Inter Simple Sequence Repeats*) utiliza um único *primer* de tamanho maior do que os usados em RAPD e com repetições. A existência de regiões de repetição em vários genomas permite que o aumento do tamanho do *primers* não incorra em pouco anelamento. Assim, resolve-se o problema da baixa especificidade de *primers* pequenos, o que aumenta a reproduzibilidade da técnica, reduzindo o efeito das variáveis nas condições de amplificação.

Outra técnica que tenta reduzir a falta de reproduzibilidade dos RAPD's e o AFLP (do inglês, *Amplified Fragment Length Polymorphism*). Neste caso, um genoma é sujeito a digestão por enzimas de restrição e os fragmentos formados ligados a adaptadores que servem de sítio de anelamento para *primers* que serão usados em uma reação de PCR. Assim, é possível amplificar fragmentos anônimos do genoma que alto grau de reproduzibilidade. As sequências que existem nos adaptadores são escolhidas arbitrariamente.

Os SNPs (do inglês, *Single Nucleotide Polymorphism*) são resultantes de alterações pontuais na sequência de nucleotídeos. Os SNPs são bialélicos, correspondendo à substituição de uma única base que ocupa uma localização em particular do genoma. Parte dos métodos de análise de SNPs depende de uma reação inicial de PCR. Os produtos desta reação podem ser analisados por diversas técnicas para diferenciação dos alelos como, por exemplo, uma eletroforese em gel.

A utilização das técnicas de PCR, diferente do que se pode pensar, também precisa assumir alguns pressupostos para que os padrões observados possam ser interpretados como variação gênica. Os mais óbvios são que o produto obtido é o

desejado. Além disso, precisa-se assumir que todos os alelos estão sendo certamente e igualmente amplificados e, por último, que os segmentos de DNA amplificados são comuns por descendência e não por convergência.

1.3.6.5. Sequenciamento

O sequenciamento é a técnica mais precisa para amostragem da variação gênica de qualquer genoma ou tipo de sequência. Sequenciamento, como diz o nome, é o processo pelo qual a ordem sequencial das partes constituintes de uma molécula de DNA ou RNA (os nucleotídeos) é determinada. A sequência é a representação linear simbólica dos nucleotídeos que, sucintamente, resume a estrutura linear do segmento da molécula de DNA (ou RNA) sequenciado.

Na década de 1970, Frederick Sanger desenvolveu o método de sequenciamento chamado de método de terminadores de cadeias (Sanger & Coulson 1975). Este método utiliza uma técnica que tem como princípio a terminação sequência-específica da reação de síntese de DNA utilizando nucleotídeos modificados, os ddNTPs (didesoxirribonucleotídeos), que são análogos aos dNTPs normais (deoxinucleotídeos). Assim, são realizadas quatro reações usando *primers* para uma região específica de DNA, *Taq* DNA polimerase, deoxinucleotídeos (A- Adenina, T- Timina, C- Citosina e G- Guanina) e um didesoxirribonucleotídeo (ddNTP de A, T, C e G) diferente para cada reação. A incorporação limitada dos terminadores na cadeia crescente de DNA interrompe a extensão da fita no ponto de inserção do nucleotídeo terminador, gerando fragmentos terminados nos 4 ddNTPs existentes. Isto produz fragmentos truncados da sequência completa de interesse. Os fragmentos obtidos, possuindo diferentes tamanhos, são separados por peso molecular em eletroforese em gel que após uma série de procedimentos permite a leitura visual da sequência. Os ddNTPs podem ser, também,

marcados com fluorocromos diferentes para cada um deles, o que permite a leitura por fluorescência, o que resulta em um método mais rápido e fácil.

Se o método de Sanger consistiu numa revolução em relação ao seqüenciamento de DNA na década de 1970, as técnicas de seqüenciamento têm sofrido, a partir da década de 1990, nova revolução com as técnicas de sequenciamento de nova geração ou NGS (*New Generation Sequencers*, do inglês). Nestas tecnologias o sequenciamento de DNA é promovido em plataformas capazes de gerar informação sobre milhões de pares de bases com grande economia de tempo e custo. Tanta eficiência advém do uso da clonagem *in vitro* e de sistemas de suporte sólido para as unidades de sequenciamento (Carvalho & Silva 2010).

Entre as NGS encontra-se as que utilizam o pirosequenciamento, desenvolvido em 1996 por Pål Nyrén e Mostafa Ronaghi. Esta técnica tem como princípio a detecção do pirofosfato liberado durante a incorporação do nucleotídeo, ao invés da terminação da cadeia de DNA com ddNTPs. De maneira geral, no pirosequenciamento, fitas simples de DNA são aneladas a *beads* (grânulos) e amplificados via EmPCR (Emulsion-based PCR). Estas moléculas de DNA associadas aos *beads* são, então, colocadas em placas, localizadas próximas a uma fibra ótica, juntamente com enzimas que produzem luz na presença de ATP. Quando nucleotídeos (dNTPs) são adicionados à reação, luzes passam a ser emitidas a medida que estes dNTPs vão sendo sequencialmente incorporados a cadeia crescente de DNA. Os diferentes sinais luminosos são registrados por um detector de fótons.

Os NGS encontram-se já na sua quarta geração (Primeira Geração-Sanger; Segunda Geração-Pirosequenciamento; Terceira Geração-*Single Molecule/Real Time*) com os seqüenciadores *Ion Torrent*. Portanto, não vale a pena se estender para além do ponto em que se foi, uma vez que em relação a estas tecnologias, tudo que se escreve já

é velho em princípio, quanto mais depois do ponto final. Contudo, o esforço de explicar princípios básicos se justifica, uma vez que qualquer salto começa pelo apoio no chão.

1.3.7. Ômicas

A genômica é uma área nova, da década de 1980. O seu desenvolvimento está ligado à modificação da forma de seqüenciamento de Frederick Sanger, ao automatizá-la e integrá-la a um sistema computadorizado (Furlan et al. 2007). Isto determinou que o processo de seqüenciamento fosse realizado em larga escala e com alto desempenho. Outro fato que, certamente, contribui muito para o estabelecimento da genômica foi o “Projeto Genoma Humano” (Lander et al. 2001; Venter et al. 2001).

A genômica pode ser estrutural ou funcional. A genômica estrutural se dedica ao sequenciamento dos nucleotídeos, mapeamento e análise de genomas, os quais são constituídos por todo o DNA, incluindo os genes estruturais, sequências regulatórias e regiões não codificantes. A genômica funcional inclui aspectos relacionados com a função do genoma em si, tais como análise de mutações e polimorfismos, bem como medida de atividade molecular. A genômica funcional é constituída pelas chamadas “-ômicas”, como a transcritômica, proteômica e a metabolômica.

A transcritômica estuda a expressão gênica, ou seja, o conjunto de transcritos de um determinado organismo, que são altamente dinâmicos e se modificam de acordo com as diferentes condições celulares e ambientais. A proteômica estuda a expressão das proteínas em larga escala, podendo ser definida pelo estudo do conjunto de proteínas expressas por um organismo, tecido ou célula, incluindo as modificações no padrão de expressão das proteínas em condições diferentes. A metabolômica visa identificar e quantificar o conjunto de metabólitos produzidos e/ou modificados por um organismo. O perfil metabólico pode fornecer um panorama geral sobre o estado fisiológico de um organismo.

A genômica funcional busca ampliar e sintetizar o conhecimento sobre genômica, transcritômica, proteômica e metabolômica no intuito de compreender as propriedades dinâmicas em níveis celulares e/ou organismos. Isso proporcionaria uma visão de como uma função biológica surge a partir da informação codificada a partir do genoma de um organismo. Existe muito conhecimento sobre a natureza e conteúdo da informação genética, assim como as tecnologias disponíveis para o sequenciamento de genomas em larga escala (Pearson & Lipman 1988). Contudo, os dados obtidos nos projetos de sequenciamento não têm contribuído da forma esperada para o melhor entendimento do papel que os genes e as proteínas desempenham no funcionamento dos organismos, uma vez que o genoma é um elemento virtualmente estático, ao passo que seus produtos, representados pelos RNAs mensageiros, proteínas e metabólitos, possuem um caráter dinâmico, caracterizado por mudanças contínuas em resposta a estímulos internos e externos.

As ômicas possibilitam posicionar as funções individuais dos genes e seus produtos dentro de um contexto global. As técnicas analíticas permitem avaliar os padrões de expressão gênica e protéica nas células e tecidos, para entender como estas macromoléculas interagem de maneira dinâmica para produzir organismos capazes de se adaptar às influências do ambiente e a situações fisiológico-metabólicas específicas. Os resultados das pesquisas em genômica funcional têm permitido um entendimento das relações gene-fenótipo, bem como dissecar as interações genótipo-ambientais que ocasionam mudanças no fenótipo dos indivíduos.

Uma abordagem bastante utilizada em genômica, particularmente em estudos com eucariontes, é o sequenciamento das Etiquetas de Sequências Expressas ou ESTs (do inglês, *Expressed Sequence Tags*), cujo objetivo é caracterizar a expressão temporal e local dos genes estruturais mediante a identificação de seus respectivos RNAs

mensageiros. Devido a grande instabilidade das moléculas de RNA, a estratégia utilizada nessa modalidade de sequenciamento é a utilização do princípio da transcrição reversa para converter o RNA mensageiro em seu DNA complementar (cDNA), que posteriormente terá as extremidades sequenciadas. Essas pequenas sequências são analisadas por programas computacionais que têm a capacidade de identificar o gene que lhes deu origem. As coleções de ESTs são muito usadas para os estudos de genômica funcional e têm sido largamente aplicadas na identificação de novos genes, mapeamento e anotação de transcritos, assim como para a construção de mapas comparativos, utilizando informações geradas em projetos conduzidos com outras espécies. Além disso, as bibliotecas de cDNA produzidas nestes projetos fornecem o substrato para a confecção das lâminas de *microarrays* utilizadas em ensaios de expressão gênica.

Microarrays é uma técnica baseada no uso de um *chip* de DNA, que é uma plataforma sólida, cuja superfície contém um grande número de genes distribuídos de forma ordenada. Embora existam variações, o tipo mais comum é aquele em que a matriz sólida é uma lâmina de vidro usada em microscopia (também chamada de slide), contendo milhares de *spots* (locais onde se encontram os genes), nos quais está depositada uma quantidade infinitamente pequena de DNA (sonda). As fontes de DNA mais comumente utilizadas como sondas são os produtos de PCR amplificados de cDNAs ou oligonucleotídeos sintetizados a partir de sequências depositadas no *GenBank*. Apesar da crescente popularidade na aplicação desta tecnologia, é necessário ressaltar que a formação do RNA mensageiro é apenas o primeiro passo de uma longa sequência de eventos que resultam na síntese das proteínas funcionais, que são, em última instância, as moléculas responsáveis pelo fenótipo das células.

Logo após a transcrição do DNA, o RNA resultante está sujeito a modificações que podem resultar na formação de diferentes isoformas de proteínas, geradas a partir de um único gene. Além disso, os mecanismos que regulam o processo de tradução, as modificações pós-traducionais e, principalmente, o processo de proteólise celular, são fatores que afetam significativamente as quantidades e as características estruturais e funcionais das proteínas. Neste sentido, o conteúdo de proteínas na célula não é um reflexo direto da abundância dos RNAs mensageiros específicos. As tecnologias que permitem a caracterização do perfil global das proteínas presentes numa célula, tecido, ou organismo, como a eletroforese bi-dimensional e a espectrometria de massa, são ferramentas usadas para comparar diferentes estados de um sistema biológico.

A aplicação conjunta das tecnologias empregadas em estudos de transcriptoma (*microarrays*) e proteoma (eletroforese bi-dimensional e espectrometria de massa) em ensaios que envolvem indivíduos geneticamente diferentes, ou submetidos a condições ambientais contrastantes, permite um entendimento das relações gene-fenótipo, e possibilita dissecar as interações genótipo-ambientais que ocasionam mudanças no fenótipo. Em Genética Marinha, as ômicas têm sido usadas, por exemplo, para a avaliação da expressão de genes em peixes expostos a elementos tóxicos. Na seção **1.6. Científico** alguns estudos com ômicas são descritos.

1.3.8. E Agora José?

A quantidade de métodos e técnicas disponíveis hoje em dia para a amostragem e estudo da variação gênica em populações de organismos marinhos (depois desta “sopa de letrinhas” é bom lembrar-se do que se trata em Genética Marinha) é imensa. Portanto, neste momento, perguntas legítimas seriam: qual técnica usar? Existe, dentre todas elas, uma que seja “A” melhor? Existe uma relação biunívoca técnica-

pergunta/problema em Genética Marinha? Obviamente não existem repostas simples para estas perguntas. O problema específico que se está investigando, geralmente, tem um papel importante na definição do método, bem como outros aspectos científicos como as características do marcador (herança biparental ou materna, dominância etc.), a forma de preservação das amostras (viva, congelada, mantida em álcool etc.) e a facilidade na sua obtenção (espécies raras, espécies de tamanho reduzido etc.). Contudo, existem alguns outros fatores que, no mundo real, também influenciam a escolha de uma ou outra técnica, por exemplo, os recursos econômicos e a tradição do laboratório e, também, o fetichismo associado às técnicas (Lewontin 1991; 1993; Nelkin & Lindee 1995; Tognolli 2003; Schlötterer 2004).

A definição do problema a ser investigado deve ser o primeiro critério de escolha. Diferentes marcadores moleculares apresentam diferentes taxas de substituição, ou seja, evoluem em velocidades diferentes. Marcadores que evoluem muito rápido são úteis para o estudo de indivíduos, famílias e populações. Isso se dá porque altas taxas de substituição determinam que a taxa de mudança, para uma dada região da molécula, seja tão rápida que indivíduos, mesmo muito aparentados, serão diferentes quando estudados por esse marcador. Por exemplo, técnicas como RAPDs produzem quantidade muito grande de variação observável. Isto é útil em estudos com organismos clonais, investigações sobre paternidade e modos de reprodução. Entretanto, como já se viu, estes marcadores são muito sensíveis as condições experimentais, o que faz com que os resultados fiquem, algumas vezes, *sub judice*. Tanto que algumas revistas não aceitam mais artigos que tenham sido conduzidos exclusivamente com RAPDs (*Molecular Ecology* é uma delas). Minisatélites de múltiplos *loci*, nestes casos, é uma alternativa, embora, o preço seja muito maior.

Marcadores que evoluem mais lentamente são mais adequados para os estudos que envolvem diferentes espécies ou táxons supra-específicos. Nesse caso, como a taxa de mudança é mais lenta, indivíduos de uma mesma população, por exemplo, não apresentarão diferenças, sendo todos iguais, quando estudados por esse marcador. No entanto, aqueles que pertencem a espécies, gêneros ou famílias diferentes apresentarão variação, o que possibilita a definição das relações de parentesco entre estes *taxa*. Neste caso, técnicas como RFLP e sequenciamento em exons ou regiões conservadas, são bons exemplos de técnica e sequências a serem usadas. Daí encontrarem-se muitos trabalhos de Filogenia com estas técnicas.

Os marcadores moleculares podem ser classificados de acordo com a existência de dominância. Técnicas como aloenzimas, microsatélites, minisatélites de *locus* único, RFLPs, SSCP e SNPs apresentam codominância, sendo homozigotos e heterozigotos facilmente identificáveis no padrão de bandas visualizado. Por outro lado, técnicas como RAPDs e minisatélites de *loci* múltiplos apresentam dominância. Neste caso, basta a presença de um único genoma que possibilite a amplificação do fragmento para que a banda seja visualizada. Assim, marcadores moleculares com situação de dominância tornam-se menos úteis para análises populacionais (Tautz 1989). Nessas casos, para se estimar as frequências gênicas a partir da interpretação do padrão de bandas, é preciso que se assuma um pressuposto muito importante e não necessariamente atendido: o Equilíbrio de Hardy-Weinberg (ver 1.4.1.2. *Testes de hipóteses*).

Outro critério importante a ser observado é o tipo de material disponível para estudo. Trabalhos com aloenzimas demandam quantidade razoável de material biológico (10 a 50mg por *locus* gênico analisado), que deve estar, obrigatoriamente, fresco ou congelado. Dessa forma, dependendo do organismo, a amostragem para os

trabalhos com aloenzimas pode ser destrutiva, impondo limitações para sua utilização em estudos com organismos muito pequenos (ou larvas), bem como com espécies raras. Trabalhos que utilizam as facilidades da técnica de PCR não oferecem esse tipo de problema. No entanto, o tempo de processamento das amostras no trabalho com aloenzimas é, tipicamente, menor do que para trabalhos com DNA. Somado a isso, o custo favorece as aloenzimas. Fica claro, portanto, que se deve fazer uma reflexão substancial no que se refere às questões relacionadas com eficiência, antes de escolher um curso de ação.

No mundo real, os laboratórios que têm muitos recursos podem usar técnicas caras e, laboratórios mais pobres, fazem bem em escolher técnicas mais baratas. Do mesmo modo, sequências já disponíveis no *GenBank* facilitam o trabalho para laboratórios que não dispõe de muitos recursos, uma vez que permitem o desenho imediato de *primers* e partir direto para a PCR. Contudo, a informação genética disponível está associada, geralmente, aos organismos mais importantes. Leia-se aqui, como importantes, aqueles organismos que têm importância econômica para quem financia a pesquisa ou que garante recursos para quem faz a pesquisa. Neste sentido, os laboratórios pobres estarão excluídos na grande maioria dos casos (o que é de se esperar, já que a ciência é uma atividade capitalista). Outro fator importante que determina a escolha da técnica, no mundo real, é a tradição, tanto de laboratórios quanto das técnicas. Uma técnica bem estabelecida deve ser a primeira opção, especialmente para os laboratórios que estão apenas começando na área. Da mesma forma, um laboratório que tenha sido pioneiro no uso de uma determinada técnica, pode ser mais conservador sobre a utilização de abordagens menos tradicionais.

As técnicas moleculares são apenas ferramentas, a utilização adequada destes instrumentos exige compreensão dos problemas que se quer abordar. Assim, a escolha

da técnica molecular deve ser baseada na eficiência, em vez de grau de sofisticação. Não é preciso amplificar ou sequenciar o DNA quando técnicas de aloenzimas podem responder à pergunta de forma mais rápida e mais barata. Este fato não tem sido claramente indicado na literatura científica, pelo contrário, tem sido observado algo perto do desdém para com o método de aloenzimas quando comparado a outros métodos, como microsatélites, por exemplo. Neste caso em específico, a desvalorização é, no mínimo, interessante, uma vez que as duas técnicas fornecem informações bastante semelhantes. Microsatélites apresentam um número maior de alelos que aloenzimas, por isso, quando as aloenzimas não apresentam variabilidade suficiente para resolver uma questão, a análise de microsatélites pode ser extremamente valiosa. Contudo, se não é este o caso, a escolha demonstra um custo, em tempo e dinheiro, que não se justifica... A não ser pelo fetichismo associado às técnicas.

Seja por modismo ou por imposições econômicas, uma tendência que têm se estabelecido em Genética Marinha (mas não só nela) é o fetichismo das técnicas, ou seja, de maneira acrítica (para falar o menos) que pesquisadores, laboratórios e periódicos científicos têm negligenciado a análise objetiva (racional!) dos problemas a serem investigados e têm simplesmente aplicado a fórmula:

**Quanto mais tubos
Quanto mais pipetagens
Quanto mais reagentes caros
Quanto mais aparelhos sofisticados**



MELHOR CIÊNCIA!!!

Não se pode falar o suficiente sobre o equívoco desta tendência que parece vir de brinde junto com a compra de qualquer produto de biologia molecular. Pesquisadores e laboratórios sérios têm sofrido com este fetichismo. Embora algumas vozes tenham se levantado contra esta tendência (ver bibliografia ao final deste capítulo, especialmente Karl e colaboradores em *Common misconceptions in molecular ecology: Echoes of the*

modern synthesis publicado em 2012 na *Molecular Ecology* (Karl et al. 2012)), ela não parece dar mostras de arrefecimento para breve. Por exemplo, numa pesquisa rápida na *Web of Knowledge*, para o ano de 2015, é possível perceber que o número de artigos científicos com RAPDs (uma das técnicas de DNA mais criticadas) é sete vezes maior que aqueles com aloenzimas (um dos marcadores moleculares mais bem estabelecidos da história). É claro que muitas interpretações podem ser dadas para isto, porém, a palavra “DNA” (um dos novos ícones da cultura de massas) associada à técnica de RAPD’s deve ter alguma coisa a ver com isso.

As premissas mais básicas subjacentes à utilização das técnicas de DNA são que os marcadores utilizados são herdáveis e independentes e que a técnica é repetível. Geralmente essas suposições são válidas para a maioria dos estudos. No entanto, ao se fazer Genética Marinha (como qualquer outra ciência) não se pode, nunca, esquecer que, por traz de todas as conclusões existem, sempre, pressupostos. Para terminar esta seção, na tabela 1.1 é apresentado um resumo daquilo que se deve ter em mente ao escolher uma técnica, método ou marcador molecular para se trabalhar com problemas de Genética Marinha.

Tabela 1.1. Resumo dos critérios que devem ser usados na escolha de um DNA alvo, tipo de sequência e técnica para o trabalho com Genética Marinha.

	CRITÉRIOS										
	Taxa de evolução	Neutralidade	Proporção do genoma	Quantidade de variação	Herança	Dominância	Número de loci	Número de indivíduos	Quantidade de material	Preservação	Custo
DNA ALVO											
<i>mtDNA</i>	Alta	1	Todo	Grande	Materna	Haplóide	Um	Muitos	Pequena	NL	4
<i>nDNA</i>	*	1	Todo	Grande	Mendeliana	Codominante	Grande	Muitos	Pequena	NL	4
<i>rDNA</i>	Baixa	1	Parcial	Grande	Depende	Codominante ²	Limitada	Muitos	Pequena	NL*	4
TIPO DE SEQUÊNCIA											
<i>Exons</i>	Baixa	1	Parcial	Pequena	Mendeliana	Codominante	Grande	Muitos	Pequena	NL	4
<i>Introns</i>	Alta	Sim	Parcial	Grande	Mendeliana	Codominante	Grande	Muitos	Pequena	NL	4
<i>VNTR</i>	Alta	Sim	Parcial	Grande	Mendeliana	Codominante	Grande	Muitos	Pequena	NL	Alto
TÉCNICAS											
<i>Aloenzimas</i>	Mediana	Discutível	Pequena	Mediana	Mendeliana	Codominante	Grande	Muitos	Razoável	L	Baixo
<i>RFLP</i>	*	1	Parcial	Mediana	Depende	Codominante	Grande ³	Muitos	Pequena	NL	Alto
<i>SSCP</i>	*	1	Parcial	Mediana	Depende	Codominante	Grande ³	Muitos	Pequena	NL	Mediano
<i>SNP</i>	*	1	Parcial	Grande	Depende	Codominante	Grande ³	Poucos	Pequena	NL	Alto
<i>RAPD</i>	*	Sim	Parcial	Grande	Mendeliana	Dominante	Grande	Muitos	Pequena	NL	Baixo
<i>Sequenciamento</i>	*	1	Todo	Grande	Depende	Codominante ²	Grande ³	Poucos	Pequena	NL	Mediano

* Depende do tipo de sequência

1- Depende da região trabalhada

2- Exceto se haplóide

3- Apenas um *locus* no caso do mtDNA

4- Depende da técnica utilizada

L- Limitado quanto à preservação, o material precisa ser fresco ou congelado

NL- Excetuando-se fixação em formol, não existe limitações quanto à preservação

Depende- no caso do nDNA, a herança é mendeliana; no caso do mtDNA, a herança é materna

1.4. CONCEITUAL

Conceitual vem de conceito, aquilo que tem base nos conceitos. Mas o que é um conceito? Este termo, conceito, como quase todos os termos chave da linguagem científica vem do latim, *conceptus* (do verbo *concipere*), e significa "coisa concebida" ou "formada na mente". Desta forma, em senso amplo conceito define tanto uma ideia quanto um juízo de valor ou opinião. No entanto, em ciência, este termo é a própria negação de um juízo de valor ou opinião. Em ciência, um conceito é aquilo que foi construído na razão para apreender os objetos do mundo. É neste sentido que se diz que os conceitos devem ser “objetivos”, ou seja, traduzir de maneira racional os objetos do mundo sem maiores (menores sempre existem) interferências de juízos de valores ou opiniões.

E o que se quer dizer aqui com conceitual? As ferramentas de trabalho de uma ciência são os conceitos com os quais interpreta e interroga o mundo de modo a obter conhecimento e, com estes, intervir de volta no mundo. Como já foi visto aqui no “Histórico” a Genética Marinha tem como avós a teoria evolutiva darwiniana e o modelo mendeliano de herança e como pais a teoria sintética da evolução e a genética molecular. Já foi visto, também, que, objetivamente falando (e aqui objetivamente já se entende como aquilo que é dito por meio dos objetos racionais criados para apreender os objetos do mundo, ou seja, os conceitos), a grande novidade da teoria evolutiva darwiniana foi a perspectiva materialista da variação e, da síntese evolutiva, entender como ocorre a mudança dos padrões de variação hereditária (genótipo e fenótipo) ao longo das gerações. Portanto, o conceitual aqui apresentado diz respeito a todos os conceitos que a Genética Marinha usa para descrever a variação presente nas populações de organismos marinhos.

1.4.1. Variação Intrapopulacional

A primeira tarefa de um evolucionista e, portanto, de um geneticista marinho é a descrição da variação intrapopulacional. Isto diz respeito a primeira e maior ruptura da teoria evolutiva darwiniana: negar a existência de um tipo (platônico) e de um sistema de causas que ordenasse a natureza (aristotélico), e estabelecer uma visada populacional para uma variação, de agora em diante, entendida como puramente contingencial.

1.4.1.1. Descrição dos dados empíricos

A variação gênica intrapopulacional pode ser quantificada através de frequências gênicas e genotípicas. A frequência de cada um dos genótipos de um *locus* (frequência genotípica) numa população pode ser obtida, facilmente, através da divisão do valor total de indivíduos encontrados para determinado genótipo pelo valor total de indivíduos amostrados na população. Por sua vez, a frequência de cada alelo (frequência gênica), neste mesmo *locus*, pode ser obtida através da divisão do valor total de um determinado alelo (em populações diploides corresponde a duas vezes o número de indivíduos homozigotos para o determinado alelo, pois cada indivíduo terá duas cópias do alelo, mais o número de indivíduos heterozigotos portadores do determinado alelo, pois cada indivíduo possuirá uma cópia do alelo) pelo valor total de alelos na população (em populações diploides corresponde a duas vezes o tamanho populacional, pois em cada indivíduo são encontrados dois alelos) (Quadro 1.4).

Quadro 1.4. Descrição matemática do cálculo das frequências genotípicas e gênicas.

A frequência genotípica pode ser descrita matematicamente como:

$$f(A_1A_1) = A_1A_1/n$$

Onde:

$f(A_1A_1)$ - frequência do genótipo A_1A_1 na população;

A_1A_1 - número de indivíduos A_1A_1 na população;

n - total de indivíduos na população.

A frequência gênica pode ser descrita matematicamente como:

$$f(A_1) = [2(A_1A_1)+(A_1A_2)]/2n$$

Onde:

$f(A_1)$ - frequência do alelo A_1 na população;

A_1A_1 - número de indivíduos A_1A_1 na população;

A_1A_2 - número de indivíduos A_1A_2 na população;

n - total de indivíduos na população.

Por exemplo, em uma população de tubarões da qual 100 indivíduos foram amostrados para um *locus* e, após análise genética, apresentou indivíduos dos seguintes genótipos:

$A_1A_1 = 23$ indivíduos, $A_2A_2 = 26$ indivíduos, $A_1A_2 = 51$ indivíduos

A frequência genotípica de A_1A_1 será:

$$f(A_1A_1) = 23/100$$

$$f(A_1A_1) = 0,23$$

E a frequência do alelo A_1 será:

$$f(A_1) = [2(23)+(51)]/200$$

$$f(A_1) = [46+51]/200$$

$$f(A_1) = 97/200$$

$$f(A_1A_1) = 0,485$$

Uma forma bastante usada para medir variação gênica é a heterozigosidade. A heterozigosidade marca a probabilidade de um indivíduo de uma determinada população ser heterozigoto em um determinado *locus*. A heterozigosidade é calculada através do número de heterozigotos em um determinado *locus* dividido pelo número total de indivíduos amostrados (Quadro 1.5).

Quadro 1.5. Descrição matemática da heterozigosidade por *locus*.

A Heterozigosidade por *locus* pode ser descrita matematicamente como:
$$h = A_1A_2/n$$

Onde:

A_1A_2 - número de indivíduos heterozigotos na população;

n - total de indivíduos na população.

Por exemplo, na população de tubarões, da qual 100 indivíduos foram amostrados para um *locus* e, após análise genética, apresentou indivíduos dos seguintes genótipos:

$A_1A_1 = 23$ indivíduos, $A_2A_2 = 26$ indivíduos, $A_1A_2 = 51$ indivíduos

A Heterozigosidade será:

$$h = 51/100$$

$$h = 0,51$$

Além da heterozigosidade apresentada no quadro 1.5 (chamada de heterozigosidade observada), existe uma outra forma de calcular a heterozigosidade nas populações naturais, que é através da heterozigosidade esperada. A heterozigosidade esperada é diferente da heterozigosidade observada porque ela é uma previsão baseada nas frequências gênicas conhecidas (Quadro 1.6).

Quadro 1.6. Descrição matemática da heterozigosidade esperada por *locus*.

A Heterozigosidade esperada por *locus* pode ser descrita matematicamente como:
$$h = 1 - \sum p_i^2$$

Onde:

h - índice de diversidade gênica

Σ - somatório;

p - é a frequência do alelo i do gene neste *locus*.

Por exemplo, na população de tubarões, da qual 100 indivíduos foram amostrados para um *locus* e, após análise genética, apresentou indivíduos dos seguintes genótipos:

$A_1A_1 = 23$ indivíduos, $A_2A_2 = 26$ indivíduos, $A_1A_2 = 51$ indivíduos

A Heterozigosidade esperada será:

$$h = 1 - [(0,485)^2 + (0,515)^2]$$

$$h = 1 - [0,235225 + 0,265225]$$

$$h = 1 - 0,50045$$

$$h = 0,49955$$

Do mesmo modo, a quantidade de variação gênica referente à amostragem de vários *loci* na mesma população pode ser apreendida a partir de medidas como o grau de polimorfismo, número de alelos e número efetivo de alelos e heterozigosidade média. O grau de polimorfismo corresponde ao número de *loci* polimórficos dividido pelo

número total de *loci* analisados. Um *locus* pode ser considerado polimórfico quando mais de um alelo for detectado. O número de alelos é, simplesmente, a média aritmética do número de alelos presente em cada *loci*. Contudo, em algumas circunstâncias, alguns *loci* podem apresentar, por exemplo, seis alelos, sendo que um deles tem uma frequência de 0,95 (ou 95%, o que é a mesma coisa) e, os cinco restantes, 0,01 (1%) cada. Neste caso, embora polimórfico, este *locus* se apresenta quase como monomórfico. Para circunscrever problemas como este, uma medida melhor que o número médio de alelos dos *loci* amostrados, é o número efetivo de alelos. O número efetivo de alelos considera a frequência de cada um dos seis alelos e é calculado com base na frequência de cada um dos alelos neste *locus*. A heterozigosidade média é calculada como a média aritmética da heterozigosidade esperada encontrada para cada um dos *loci* (Quadro 1.7).

Quadro 1.7. Descrição matemática de nível de polimorfismo, número médio de alelos, número efetivo de alelos e heterozigosidade média.

O grau de polimorfismo pode ser descrito matematicamente como:

$$P = np/nl$$

Onde:

P- grau de polimorfismo;

np- número de *loci* polimórficos;

nl- número de *loci* total.

O número médio de alelos pode ser descrito matematicamente como:

$$Nal = \Sigma na/nl$$

Onde:

Nal- número médio de alelos;

Σ - somatório;

na- número de alelos em cada *locus*;

nl- número de *loci*.

O número efetivo de alelos pode ser descrito matematicamente como:

$$Ne = 1/(\sum p_i^2)$$

Onde:

Ne- número efetivo de alelos;

Σ - somatório;

p- frequência do alelo i do gene neste *locus*.

A Heterozigosidade média pode ser descrita matematicamente como:

$$H = \Sigma(h/nl)$$

Onde:

Σ - somatório;

nl- número de *loci*;

h- heterozigosidade esperada por *locus*.

Por exemplo, em uma população de anêmonas-do-mar foram amostrados três *loci* que apresentaram as seguintes frequências gênicas:

Locus A: $A_1 = 0,063; A_2 = 0,874; A_3 = 0,063$

Locus B: $B_1 = 0,325; B_2 = 0,650; B_3 = 0,025$

Locus C: $C_1 = 1,000$

Então, o grau de polimorfismo será:

$$P = 2/3$$

$$P = 0,66$$

O número médio de alelos será:

$$Nal = (3+3+1)/3$$

$$Nal = 2,33$$

O número efetivo de alelos para o *locus A* será:

$$Ne = 1/[(0,063)^2 + (0,874)^2 + (0,063)^2]$$

$$Ne = 1/0,7718$$

$$Ne = 1,29$$

A Heterozigosidade média será:

$$H = (0,2282 + 0,4714 + 0)/3$$

$$H = 0,2332$$

Até este momento, a descrição da variação intrapopulacional tem sido feita de forma aritmética e, mais que isso, com forte influência dos dados empíricos, ou seja, depois de amostrados os genótipos dos indivíduos por meio de uma das técnicas (e métodos) descritas em 1.3. Isto sempre pode ser feito, uma vez que os genótipos pudessem ser amostrados (mesmo que fenotipicamente). Contudo, uma questão muito importante em Genética Marinha é como esta variação que foi descrita vai se comportar ao longo das gerações. Dito de outra forma, uma vez que a variação gênica de uma população natural de organismos marinhos foi descrita, como ela vai se comportar na próxima geração?

1.4.1.2. Testes de hipóteses

A questão de como saber qual o comportamento das frequências genotípicas ao longo das gerações foi respondida em 1908, e não foi um geneticista quem o fez! Foram um matemático inglês, Godfrey Hardy e um médico alemão, Wilhelm Weinberg, que propuseram uma explicação matemática simples de como as freqüências gênicas (ou alélicas) e genotípicas de uma população se comportariam ao longo das gerações, a famosa Lei de Hardy-Weinberg (Hardy 1908; 1967; Stern 1943; Crow 1987; 1999).

O que Hardy e Weinberg definiram foi que, dadas as frequências gênicas, é possível inferir as frequências genotípicas esperadas de uma população diplóide de reprodução sexuada. De acordo com eles, as frequências genotípicas esperadas são dadas pela soma dos quadrados das frequências alélicas. Assim, as frequências genotípicas dependem somente das frequências gênicas dos genitores e não de suas frequências genotípicas. Do mesmo modo, a partir das frequências genotípicas observadas, usando a mesma formulação, é possível estimar as frequências gênicas. No

quadro 1.8 esta interpretação da relação entre frequências gênicas e genotípicas pela Lei de Hardy Weinberg está descrita, tanto algébrica quanto aritméticamente.

Quadro 1.8. Descrição matemática da Lei de Hardy-Weinberg.

As frequências genotípicas esperadas podem ser descritas algébricamente como:
$$(p+q)^2 = p^2 + 2pq + q^2 = 1$$

Onde:

- p- frequência observada do alelo A₁ na população;
- q- frequência observada do alelo A₂ na população;
- p²- frequência esperada do genótipo A₁A₁ na população;
- 2pq- frequência esperada do genótipo A₁A₂ na população;
- q²- frequência esperada do genótipo A₂A₂ na população.

Retomando o exemplo do quadro 1.4 no qual, de uma população de tubarões, foram amostrados 100 indivíduos para um locus e encontrados os genótipos:

$$A_1A_1 = 23 \text{ indivíduos}, A_2A_2 = 26 \text{ indivíduos}, A_1A_2 = 51 \text{ indivíduos}$$

As frequências genotípicas observadas são:

$$\begin{aligned}f(A_1A_1) &= 0,23 \\f(A_1A_2) &= 0,51 \\f(A_2A_2) &= 0,26\end{aligned}$$

As frequências gênicas são:

$$\begin{aligned}f(A_1) &= 0,485 = p \\f(A_2) &= 0,515 = q\end{aligned}$$

Portanto, as frequências genotípicas esperadas serão:

$$\begin{aligned}p^2 &= (0,485)^2 = 0,235 \\2pq &= 2 \times 0,485 \times 0,515 = 0,500 \\q^2 &= (0,515)^2 = 0,265\end{aligned}$$

Esta relação matemática entre frequências gênicas observadas e frequências genotípicas esperadas definem o comportamento dos genótipos na população ao longo de todas as gerações se, e somente se, a população estiver em equilíbrio, de agora em diante, chamado de equilíbrio de Hardy-Weinberg. Mas o que é o equilíbrio de Hardy-Weinberg? O equilíbrio de Hardy-Weinberg é uma população repetir as suas frequências gênicas e genotípicas ao longo de todas as gerações. E como uma população pode se manter em equilíbrio de Hardy-Weinberg? Se, e somente se, nada estiver acontecendo com ela! Dito de outra forma, o equilíbrio de Hardy-Weinberg afirma que nada muda numa população se nada acontece com ela. E o que pode acontecer na vida de uma população? Muita coisa:

- 1) Nesta população diplóide de reprodução sexuada pode ser que a chance de encontros não seja igual para todos;
- 2) Pode ser, também, que nem todos tenham a mesma competência para se reproduzir;
- 3) Indivíduos de outras populações podem estar interferindo na reprodução dos indivíduos desta população.

E assim por diante. Na verdade, para que uma população consiga se manter em equilíbrio de Hardy-Weinberg ela precisa atender a muitas condições.

De um ponto de vista matemático, o que se diz é que a Lei de Hardy-Weinberg só funciona se alguns pressupostos forem mantidos. Estes pressupostos são:

- 1) Panmixia ou ausência de cruzamentos preferenciais;
- 2) Ausência de Seleção Natural;
- 3) Ausência de Deriva Genética;
- 4) Ausência de Migração;
- 5) Ausência de Mutação.

Ou seja, se nenhuma força evolutiva estiver atuando na população. Na prática, o que é feito em Genética Marinha é, a partir de amostras das populações naturais, inferir as frequências gênicas e, a partir delas, calcular as freqüências genotípicas esperadas pela Lei de Hardy-Weinberg (como foi visto no Quadro 1.8). A partir daí, as frequências genotípicas observadas e esperadas podem ser comparadas. Estas comparações servem para se testar se a Lei de Hardy-Weinberg está sendo seguida. Neste ponto, o trabalho em Genética Marinha sai da pura e simples descrição da variação gênica intrapopulacional e começa a adentrar o campo do teste de hipóteses. Mas o que se quer testar? Se é possível ou não rejeitar a hipótese de equilíbrio de Hardy-Weinberg para uma dada população a partir da amostra que foi tomada. Isto é feito a partir dos desvios

entre o observado e o esperado pela Lei de Hardy-Weinberg para as freqüências genotípicas.

O desvio entre as frequências observadas e esperadas pode ser usado como um importante indicador da dinâmica populacional. Como a Lei de Hardy-Weinberg assume como pressupostos a existência de panmixia (todos os cruzamentos são realizados ao acaso), ausência de seleção natural, deriva genética, migração e mutação atuando na população; quando, na amostra de uma população, os valores observados são muito diferentes daqueles esperados, diz-se que não é possível aceitar a hipótese de equilíbrio. Neste caso, uma ou mais dos pressupostos não está sendo respeitado. Contudo, existem desvios pequenos e desvios grandes. Quanto pequeno pode ser um desvio para que ele seja negligenciável? Quando um desvio pequeno deixa ser pequeno e deve ser considerado? O problema do tamanho dos desvios é aquilo que se define como significância estatística.

Para compreender a significância estatística de um desvio pode-se pensar em um exemplo concreto e simples: um jogo de moedas. Imagine que uma moeda seja jogada para cima dez vezes. A chance de se obter cara e coroa em dez jogadas é a mesma, contudo, é possível se encontrar desvios em relação aos valores esperados de cinco caras e cinco coroas, apenas por acaso. Se esta operação (jogar uma moeda 10 vezes) for repetida, por exemplo, 500 vezes, ao final será possível saber a frequência (quantas vezes em 500) cada um dos desvios ocorreu. Para ilustrar o que está se tentando explicar aqui, isto foi feito (com toda paciência do mundo) e os resultados estão mostrados na figura 1.18.

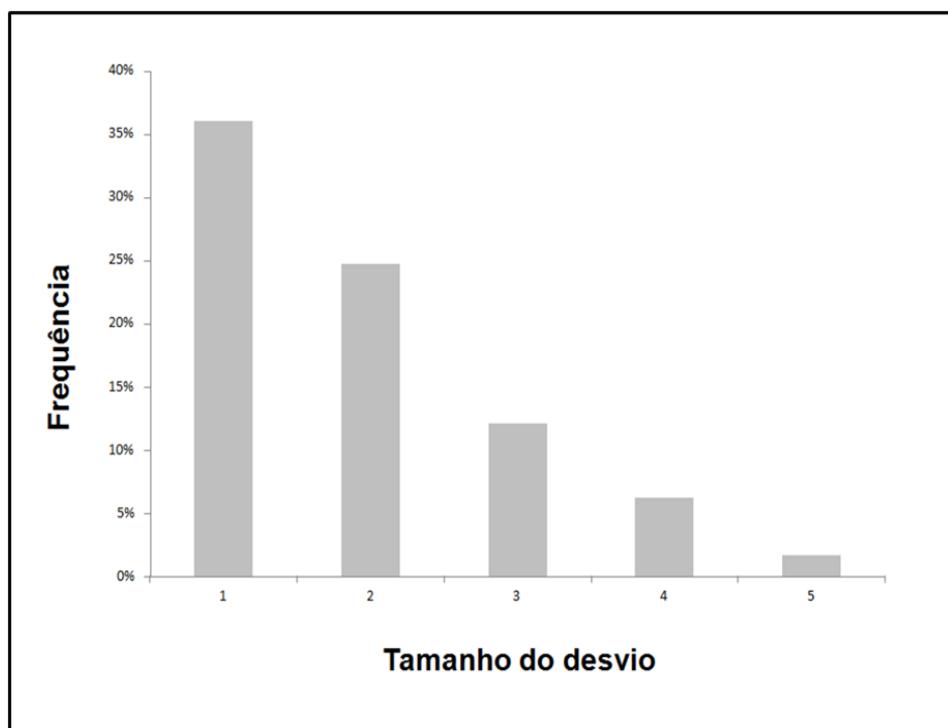


Figura 1.18. Histograma de distribuição da frequências de cada um dos desvios da proporção esperada de cinco caras e cinco coroas em dez jogadas de uma moeda. A operação foi repetida 500 vezes. Os cinco desvios são: 1- quatro caras e seis coroas (ou vice-versa); 2- três caras e sete coroas (ou vice-versa); 3- duas caras e oito coroas (ou vice-versa); 4- uma cara e nove coroas (ou vice-versa) e 5- dez caras (ou 10 coroas).

Estudando a figura 1.18 é possível observar que os desvios menores (o evento de caírem seis caras para quatro coroas, por exemplo), ocorrem muito mais vezes que os desvios maiores. O evento de caírem 10 caras, por exemplo, tem a menor frequência de todas. Desta forma, a significância estatística de um desvio está sempre associada à chance de observá-lo apenas por acaso. Quando os desvios têm uma chance muito pequena de ocorrer, admite-se que ele não deve estar acontecendo por acaso, mas que, alguma coisa, deve estar determinando o desvio. Mas qual deve ser a chance considerada pequena o suficiente para se considerar que o desvio não está sendo determinado pelo acaso? Em ciência este número mágico é, geralmente, 5% das vezes. Ou seja, toda vez que um desvio tiver uma chance de ocorrência menor que 5% das vezes, rejeita-se a hipótese de que ele está sendo determinado pelo acaso. Logo, é

provável que alguma coisa esteja acontecendo. Obviamente não é preciso se jogar tantas moedas, as expectativas e os desvios podem ser encontrados em tabelas.

Apenas mais um pequeno detalhe, toda vez que se joga uma moeda 10 vezes e se obtém 6 caras, sabe-se diretamente dos dados que caíram 4 coroas. Se caíram 8 caras, sabe-se que o número de coroas foi 2 e assim por diante. Ou seja, dado o valor de um evento, o outro não pode variar. A isto se dá o nome de graus de liberdade. Em um jogo de moedas o número de graus de liberdade é 1, isto porque existem dois eventos e o número de graus de liberdade é sempre igual ao número de eventos menos 1. Portanto, em um jogo de dados com seis lados, os graus de liberdade serão 5 ($\text{graus de liberdade} = 6 - 1 = 5$). Tudo que foi dito aqui pode ser conferido no quadro 1.9, no qual a hipótese de Hardy-Weinberg é testada para uma amostra. O teste em questão chama-se qui-quadrado e pode ser obtido dos resultados de variação gênica de marcadores codominantes como aloenzimas e microsatélites.

Quadro 1.9. Teste da hipótese de Hardy-Weinberg por qui-quadrado.

O primeiro passo para se testar a hipótese de equilíbrio Hardy-Weinberg é calcular os desvios entre as frequências genotípicas observadas e esperadas. Sabe-se, pelo quadro 1.4, que uma amostra de uma população de tubarões apresentou indivíduos dos seguintes genótipos:

$$A_1 A_1 = p^2 = 23 \text{ indivíduos}$$

$$A_1 A_2 = 2pq = 51 \text{ indivíduos}$$

$$A_2 A_2 = q^2 = 26 \text{ indivíduos}$$

No quadro 1.8 foram calculadas as frequências genotípicas esperadas para esta amostra:

$$p^2 = 0,235$$

$$2pq = 0,500$$

$$q^2 = 0,265$$

Com os dados do quadro 1.8 é possível calcular os valores esperados para cada um dos genótipos numa amostra de 100 indivíduos. Assim:

$$p^2 = 0,235 \times 100 = 23,5$$

$$2pq = 0,500 \times 100 = 50$$

$$q^2 = 0,265 \times 100 = 26,5$$

Agora, é possível se calcular os desvios usando a fórmula do qui-quadrado:

$$\chi^2 = \sum [(o - e)^2 / e]$$

Onde:

χ^2 - qui-quadrado;

o - frequência genotípica observada;

e - frequência genotípica esperada;

gl - grau de liberdade.

Portanto, os desvios serão:

$$\text{Para o genótipo } A_1 A_1 = [(23-23,5)^2/23,5] = [(-0,5)^2/23,5] = [0,25/23,5] = 0,01$$

$$\text{Para o genótipo } A_1 A_2 = [(51-50)^2/50] = [(1,0)^2/50] = [1/50] = 0,02$$

$$\text{Para o genótipo } A_2 A_2 = [(26-26,5)^2/26,5] = [(-0,5)^2/26,5] = [0,25/26,5] = 0,01$$

O qui-quadrado será, então:

$$\chi^2 = (0,01 + 0,02 + 0,01) = 0,04$$

Como são três genótipos e dois alelos, os graus de liberdade, neste caso, serão calculados como:

$gl = \text{genótipos} - \text{alelos}$

Neste caso em específico:

$$gl = 3-2 = 1$$

Com os valores de qui-quadrado ($\chi^2 = 0,04$) e graus de liberdade ($gl = 1$) é possível, utilizando-se uma tabela de qui-quadrado, saber qual a chance (ou probabilidade) de se obter um desvio desta magnitude.

G.L.\P	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,01
1	0,016	0,064	0,148	0,455	1,074	1,642	2,706	3,841	6,635
2	0,211	0,446	0,713	1,386	2,408	3,219	4,605	5,991	9,210
3	0,584	1,005	1,424	2,366	3,665	4,642	6,251	7,815	11,345
4	1,064	1,649	2,195	3,357	4,878	5,989	7,779	9,488	13,277
5	1,610	2,343	3,000	4,351	6,064	7,289	9,236	11,07	15,080

Que, neste caso, é maior que 70%, portanto, não é possível rejeitar a hipótese de Hardy-Weinberg para esta amostra.

Os cálculos do quadro 1.9 servem para ilustrar como a hipótese de Hardy-Weinberg é testada, mas, na prática, todos estes cálculos são realizados diretamente dos dados a partir de *softwares*.

1.4.1.3. Em nível de sequências

Sequências de DNA informam o máximo possível sobre a variação gênica, pois cada nucleotídeo em um grupo de sequências alinhadas pode ser considerado individualmente. Assim, o sequenciamento revela quais são os nucleotídeos em um conjunto de pares de bases, o que é definido como um haplótipo (Figura 1.19). As diferenças de nucleotídeos determinam a diversidade de haplótipos.

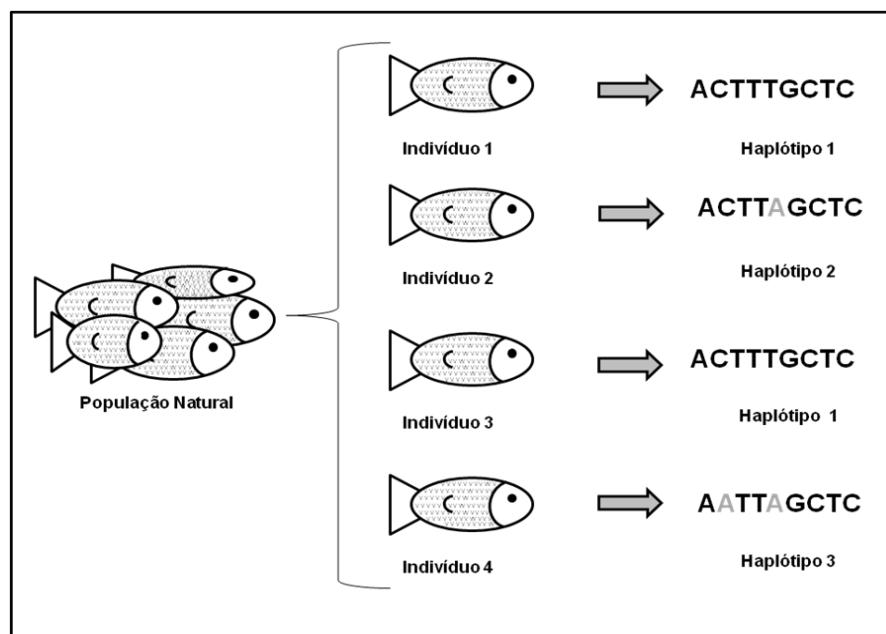


Figura 1.19. Representação dos haplótipos de quatro indivíduos de uma população natural de peixes. Observe que a sequência de nucleotídeos do indivíduo um é a mesma sequência de nucleotídeos do indivíduo três, isso significa que o haplótipo um é igual ao haplótipo três.

A variação gênica nas sequências de DNA pode ser descrita em relação ao número de sítios de nucleotídeos que são polimórficos em uma amostra de sequências (os haplótipos descritos acima) e, também, em função do número de sítios de

nucleotídeos que são heterozigotos (apresentam diferentes nucleotídeos na mesma posição entre diferentes sequências). Em um grupo de sequências alinhadas, o número de sítios segregantes (sítios de nucleotídeos que são polimórficos na amostra) é definido como “S”, ou seja, o número de sítios de nucleotídeos que diferem entre as sequências alinhadas (Quadro 1.10).

Quadro 1.10. Número de sítios segregantes.

Número de sítios segregantes é simbolizado por “S”, onde: S = Número de sítios de nucleotídeos que diferem entre as sequências alinhadas.
Retomando o exemplo da figura 1.19 no qual, de uma população de peixes, foram amostrados 4 indivíduos e encontrados os haplótipos:
Haplótipo do indivíduo 1 = ACTTGCTC
Haplótipo do indivíduo 2 = ACTTAGCTC
Haplótipo do indivíduo 3 = ACTTGCTC
Haplótipo do indivíduo 4 = AATTAGCTC
O número de sítios segregantes é:
S = 2

A medida mais comum para a variação gênica, contudo, é a heterozigosidade (H) que, em se tratando de sequências, se denomina diversidade haplotípica (h). A diversidade haplotípica mede a probabilidade de, retirando ao acaso dois haplótipos de uma amostra, eles serem diferentes (Quadro 1.11).

O número de nucleotídeos que diferem em qualquer par de sequências alinhadas são as diferenças de nucleotídeos. O número médio de diferenças de nucleotídeos para cada comparação par a par possível entre as sequências alinhadas é a diversidade nucleotídica (p) das sequências comparadas. O valor de p para qualquer amostra é, portanto, definido como número total de diferenças de nucleotídeos dividido pelo número total de comparações par a par (Quadro 1.12).

A descrição da variação gênica feita aqui em nível de sequências também foi feita de forma simples, aritmética e definida pelos dados empíricos, ou seja, depois de obtidas as sequências pelo método de Sanger, por exemplo. Também aqui, espera-se

que as sequências sigam a Lei de Hardy-Weinberg. Contudo, de posse de sequências, testes de hipóteses um pouco mais sofisticados podem ser realizados e isto será visto adiante nas seções referentes ao relógio molecular (**I.4.3**) e coalescência (**I.4.4**)

Quadro 1.11. Diversidade haplotípica.

A diversidade haplotípica pode ser definida matematicamente como:

$$h = \Sigma(1 - f_{xi}^2)/NL$$

Onde:

f_{xi} - freqüência do haplótipo i no loco x;

NL- número de *loci*.

Retomando o exemplo da figura 1.19 no qual, de uma população de peixes, foram amostrados 4 indivíduos e encontrados os haplótipos:

Haplótipo do indivíduo 1 = ACTTGCTC

Haplótipo do indivíduo 2 = ACTTAGCTC

Haplótipo do indivíduo 3 = ACTTGCTC

Haplótipo do indivíduo 4 = AATTAGCTC

As frequências haplotípicas são:

f Haplótipo 1 = $2/4 = 0,5$

f Haplótipo 2 = $1/4 = 0,25$

f Haplótipo 3 = $1/4 = 0,25$

E a diversidade haplotípica:

$$h = [1 - (0,5^2 + 0,25^2 + 0,25^2)]/1$$

$$h = 1 - (0,25 + 0,0625 + 0,0625)$$

$$h = 1 - 0,375$$

$$h = 0,625$$

Quadro 1.12. Diversidade nucleotídica.

A diversidade nucleotídica pode ser definida matematicamente como:

$$p = (\Sigma(\Sigma f_x f_y p_{xy})) / N_{hapl}$$

Onde:

f_x - freqüência do haplótipo x;

f_y - freqüência do haplótipo y;

p_{xy} - distância p entre os dois haplótipos;

N_{hapl} - número de haplótipos.

Retomando o exemplo da figura 1.19 no qual, de uma população de peixes, foram amostrados 4 indivíduos e encontrados os haplótipos:

Haplótipo do indivíduo 1 = ACTTGCTC

Haplótipo do indivíduo 2 = ACTTAGCTC

Haplótipo do indivíduo 3 = ACTTGCTC

Haplótipo do indivíduo 4 = AATTAGCTC

Sendo as frequências haplotípicas:

f Haplótipo 1 = 0,5

f Haplótipo 2 = 0,25

f Haplótipo 3 = 0,25

E as diferenças entre os haplótipos:

Haplótipo 1 = ACTTGCTC

Haplótipo 2 = ACTTAGCTC

Possuem uma diferença no total de nove nucleotídeos = 1/9 = 0,11

	<i>Haplótipo 2</i>	<i>Haplótipo 3</i>
<i>Haplótipo 1</i>	0,11	0,22
<i>Haplótipo 2</i>	***	0,11

A diversidade nucleotídica:

$$p = [(0,5 \times 0,25 \times 0,11) + (0,5 \times 0,25 \times 0,22) + (0,25 \times 0,25 \times 0,11)] / 3$$

$$p = [0,014 + 0,028 + 0,007] / 3$$

$$p = 0,048 / 3$$

$$p = 0,016$$

1.4.2. Variação Interpopulacional

A segunda tarefa de um geneticista marinho, portanto, de um evolucionista é a descrição da variação interpopulacional. Isto diz respeito ao corolário fundamental da perspectiva materialista da variação estabelecida por Darwin: entender a especiação como um processo de fracionamento da variação intrapopulacional em variação interpopulacional, tanto no espaço (geografia) quanto no tempo (gerações). Uma interpretação das diferenças populacionais e taxonomicas revolucionária e, até hoje, mal compreendida, contudo, com muitas maneiras de ser medida.

1.4.2.1. Estrutura populacional

A extensão da variação observada interpopulacionalmente resulta de um balanço de forças que produzem uma variação gênica local e forças que tendem a produzir uma homogeneidade gênica. A mutação e a deriva gênica (em populações de tamanho finito) e a seleção natural (que favorece as adaptações às condições ambientais locais) são exemplos de forças evolutivas que levam à diferenciação das populações. Já migração, favorecendo o chamado fluxo gênico, se opõe à diferenciação. O balanço entre essas forças determina a escala e o padrão de divergência entre as populações.

A dispersão de indivíduos ou gametas, com posterior participação na reprodução (migração) é um dos principais fenômenos que exercem influência na estruturação de organismos marinhos. O fluxo gênico, proveniente dos processos migratórios, promove a evolução quando espalha novas combinações de genes ou, freia a evolução, quando impede adaptações locais. Apesar da aparente homogeneidade nos ambientes marinhos, populações de locais diferentes podem se diferenciar em diferentes escalas devido à existência de barreiras geográficas e seleção após assentamento (Provine 1971; 1989; Wright 1978; Wright & Provine 1986).

A forma mais antiga e amplamente utilizada para medir a diferenciação entre populações é através dos cálculos dos índices de fixação de Wright (ou estatística F), a partir de estimativas de heterozigosidade. Estes índices são: F_{IS} , F_{ST} e F_{IT} e se baseiam na variabilidade dentro e entre as subpopulações (Wright 1951). O F_{IS} mede o grau de endocruzamento dentro de cada subpopulação, sendo uma boa indicação da deficiência de heterozigotos na subpopulação (Figura 1.20, Quadro 1.13). O índice de fixação, F_{ST} , por sua vez, mede o grau de endocruzamento de cada subpopulação em relação à população total (Figura 1.21, Quadro 1.14). O F_{IT} é o índice de endocruzamento total e inclui tanto a contribuição devido ao cruzamento não panmítico (cruzamento preferencial dentro de cada subpopulação, F_{IS}), como o efeito da divisão em subpopulações (F_{ST}) (Figura 1.22, Quadro 1.15).

Quadro 1.13. Estimativa do F_{IS} .

O nível de endocruzamento dentro de uma subpopulação devido aos cruzamentos preferenciais pode ser calculado por:

$$F_{IS} = (H_E - H_O)/H_E$$

Onde:

H_E - é a heterozigosidade esperada;

H_O - é a heterozigosidade observada.

Por exemplo, tendo sido amostradas três populações de *Coscinasterias tenuispina* (uma espécie de estrela do mar) e estimada a variação gênica intrapopulacional para um locus como:

População 1:

$A_1A_1 = 53$ indivíduos, $A_2A_2 = 13$ indivíduos, $A_1A_2 = 34$ indivíduos

População 2:

$A_1A_1 = 69$ indivíduos, $A_2A_2 = 9$ indivíduos, $A_1A_2 = 22$ indivíduos

População 3:

$A_1A_1 = 30$ indivíduos, $A_2A_2 = 30$ indivíduos, $A_1A_2 = 40$ indivíduos

As heterozigosidades observadas serão (rever Quadro 1.5 na seção 1.4.1.1. Descrição dos dados empíricos):

População 1:

$$H_O = 34 / 100 = 0,34$$

População 2:

$$H_O = 22 / 100 = 0,22$$

População 3:

$$H_O = 40 / 100 = 0,40$$

As heterozigosidades esperadas serão (rever Quadro 1.6 na seção 1.4.1.1. Descrição dos dados empíricos):

População 1:

$$H_E = 1 - [(0,70)^2 + (0,30)^2] = 1 - [0,49 + 0,09] = 1 - 0,58 = 0,42$$

População 2:

$$H_E = 1 - [(0,80)^2 + (0,20)^2] = 1 - [0,64 + 0,04] = 1 - 0,68 = 0,32$$

População 3:

$$H_E = 1 - [(0,50)^2 + (0,50)^2] = 1 - [0,25 + 0,25] = 1 - 0,50 = 0,50$$

Portanto, o F_{IS} para cada uma das três populações será:

População 1:

$$F_{IS} = (0,42 - 0,34) / 0,42 = 0,08 / 0,42 = 0,19$$

População 2:

$$F_{IS} = (0,32 - 0,22) / 0,42 = 0,10 / 0,32 = 0,31$$

População 3:

$$F_{IS} = (0,50 - 0,40) / 0,50 = 0,10 / 0,50 = 0,20$$

O F_{IS} entre as três populações será:

$$F_{IS} = (0,19 + 0,31 + 0,20) / 3$$

$$F_{IS} = 0,70 / 3$$

$$F_{IS} = 0,23$$

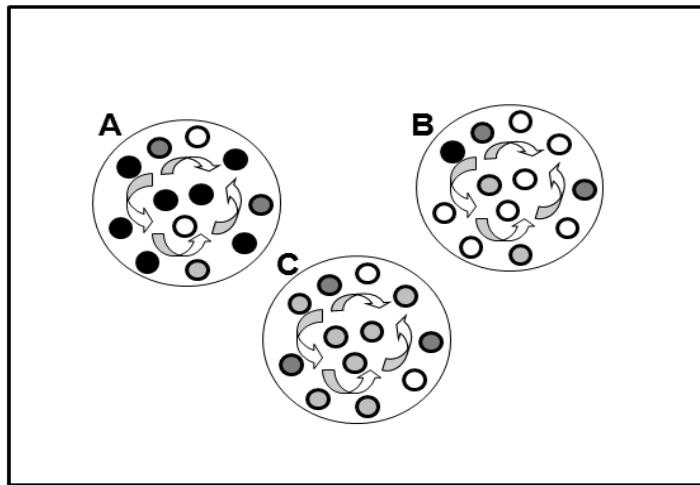


Figura 1.20. Representação gráfica do F_{IS} que mede o índice de endocruzamento dentro de cada subpopulação. Ou seja, existe uma chance maior de reprodução entre os indivíduos dentro de cada subpopulação (A, B e C) do que entre os indivíduos da população A com aqueles das populações B e C; uma maior chance de cruzamento entre os indivíduos de dentro da população B do que com aqueles das populações A e C e assim por diante.

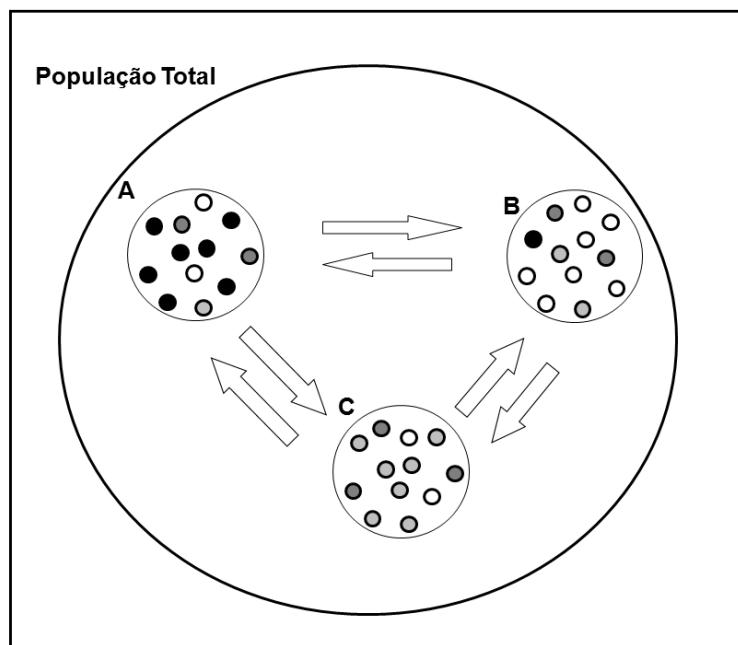


Figura 1.21. Representação gráfica do F_{ST} que mede o índice de endocruzamento de cada subpopulação em relação à população total. Ou seja, quanto da variação total está entre as subpopulações e não dentro das subpopulações.

Quadro 1.14. Representação matemática do F_{ST} .

O endocruzamento de cada subpopulação em relação à população total pode ser medido por:

$$F_{ST} = Vp/(p(1-p))$$

Onde:

p- é a média das frequências do alelo p das subpopulações;

Vp- é a variância das frequências do alelo p das subpopulações.

Para o exemplo do quadro 1.13, frequências gênicas serão (rever Quadro 1.4 na seção 1.4.1.1. Descrição dos dados empíricos):

População 1:

Locus A: $A_1 = 0,70$; $A_2 = 0,30$

População 2:

Locus A: $A_1 = 0,80$; $A_2 = 0,20$

População 3:

Locus A: $A_1 = 0,50$; $A_2 = 0,50$

O F_{ST} de cada alelo entre as três populações será:

Alelo 1

Média das frequências do alelo 1 entre as subpopulações (rever Quadro 1.1 na seção 1.3) será:
 $(0,7 + 0,8 + 0,5) / 3 = 0,67$

Variância das freqüências do alelo 1 entre as subpopulações (rever Quadro 1.2 na seção 1.3) será:

$$[(0,7-0,67)^2 + (0,8-0,67)^2 + (0,5-0,67)^2]/(3-1) = [0,001+0,018+0,028]/2 = [0,047]/2 = 0,023$$

Portanto, o F_{ST} para o alelo 1 entre as subpopulações será:

$$F_{ST} (\text{alelo } 1) = 0,023/[0,67 \times (1 - 0,67)] = 0,023/[0,67 \times 0,33] = 0,023/[0,221] = 0,105$$

Alelo 2

$$\text{Média} = (0,3 + 0,2 + 0,5) / 3 = 0,33$$

$$\text{Variância} = [(0,3-0,33)^2 + (0,2-0,33)^2 + (0,5-0,33)^2]/(3-1) = [0,047]/2 = 0,023$$

$$F_{ST} (\text{alelo } 2) = 0,023/[0,33 \times (1 - 0,33)] = 0,023/[0,221] = 0,105$$

O F_{ST} deste locus para as três subpopulações será:

$$F_{ST} = [F_{ST} (\text{alelo } 1) + F_{ST} (\text{alelo } 2)]/2$$

$$F_{ST} = (0,105 + 0,105) / 2$$

$$F_{ST} = 0,105$$

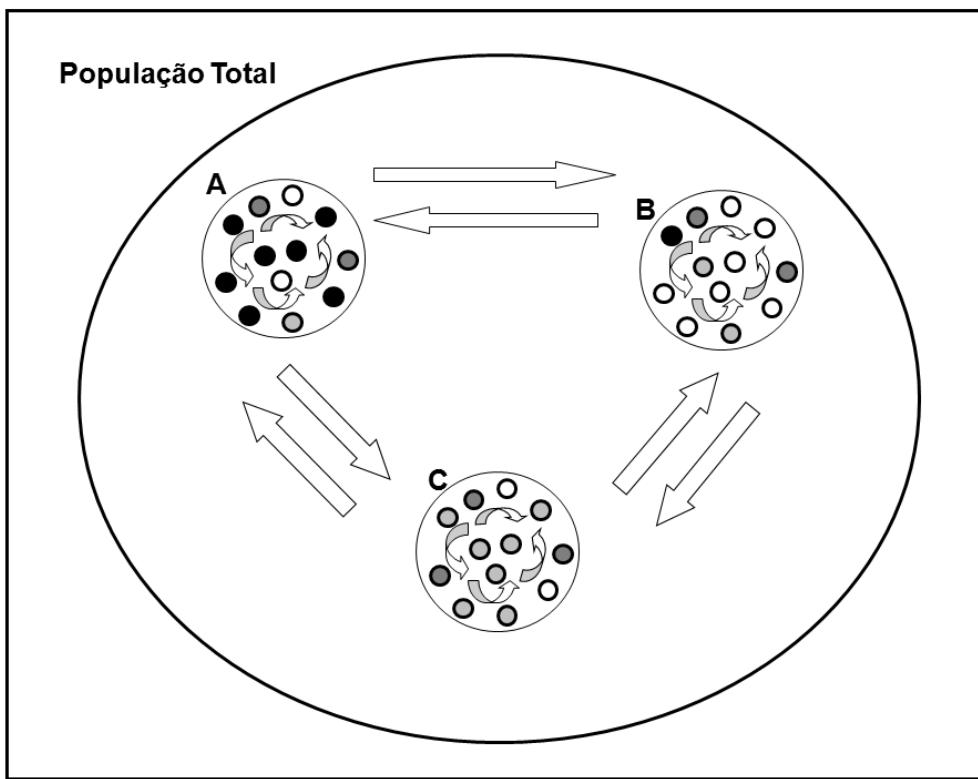


Figura 1.22. Representação gráfica do F_{IT} que mede o índice de endocruzamento total que inclui tanto os efeitos do cruzamento preferencial dentro de cada subpopulação (F_{IS}), quanto o efeito da subdivisão populacional (F_{ST}).

Quadro 1.15. Representação matemática do F_{IT} .

Os índices de fixação de Wright estão relacionados da seguinte forma:

$$1 - F_{IT} = (1 - F_{IS}) (1 - F_{ST})$$

Onde:

F_{IT} - índice de endocruzamento total;

F_{IS} - índice de endocruzamento dentro de cada subpopulação;

F_{ST} - índice de endocruzamento de cada subpopulação em relação à população total.

Para o exemplo dos quadros 1.13 e 1.14 têm-se:

$$F_{IS} = 0,230$$

$$F_{ST} = 0,105$$

O F_{IT} entre as três populações será:

$$1 - F_{IT} = (1 - 0,230) (1 - 0,105)$$

$$1 - F_{IT} = (0,770) (0,895)$$

$$1 - F_{IT} = 0,689$$

$$- F_{IT} = 0,689 - 1$$

$$- F_{IT} = - 0,311 \times (-1)$$

$$F_{IT} = 0,311$$

O F_{ST} está diretamente relacionado com a variância das frequências alélicas entre as populações. Se o valor de F_{ST} é baixo, quer dizer que a frequência dos alelos dentro da cada subpopulação é similar; se seu valor for alto, significa que existem diferenças razoáveis entre as freqüências dos alelos nas diferentes subpopulações. O valor de F_{ST} varia de zero a um. $F_{ST} = 1$ indica subpopulações com a fixação de alelos distintos entre elas. $F_{ST} = 0$ indica subpopulações com frequências alélicas idênticas, ou seja, não existe diferença genética entre as subpopulações. De uma maneira geral, é possível dizer que valores de F_{ST} entre 0 e 0,05 indicam pequena diferenciação genética. Entre 0,05 e 0,15 o F_{ST} indica diferenciação genética moderada. F_{ST} entre 0,15 e 0,25 indica grande diferenciação genética. Acima de 0,25 a diferenciação genética é muito grande.

Na literatura de Genética Marinha abundam diferentes coeficientes e índices de fixação, pois várias estatísticas análogas aos índices de estruturação de Wright foram surgindo ao longo dos anos. Estes diferentes métodos levam em conta que o processo de geração de variabilidade é diferente em diferentes circunstâncias e para diferentes tipos de marcadores moleculares. A maioria dos estimadores baseados em F_{ST} foi desenvolvida para modelos de populações simétricas. Mastoshi Nei desenvolveu o G_{ST} , que consiste em um parâmetro análogo ao F_{ST} com base em medidas de heterozigosidade. Pela metodologia de Nei, a diversidade genética de qualquer população pode ser decomposta em termos de heterozigosidade dentro e entre subpopulações. A grande contribuição de Nei foi estender o parâmetro F_{ST} de Wright para um contexto de análise de dados que leva em consideração múltiplos alelos num determinado *locus*, em vez de apenas dois como no modelo de Wright (Nei 1972; 1978). Bruce Weir e Clark Cockerham propuseram uma série de parâmetros populacionais análogos aos índices de fixação de Wright (f , F e θ) nos quais a estimativa é feita a

partir de componentes de variância obtidos da análise de variância das frequências alélicas nos diversos níveis de hierarquia populacional e, levando em consideração, diferentes tamanhos de populações (Weir & Cockerham 1984). Montgomery Slatkin desenvolveu o R_{ST} que é um estimador apropriado para medir a subestruturação populacional quando os dados genéticos são obtidos a partir de *loci* de microssatélites. Neste caso, assume-se que estes *loci* sofrem mutação seguindo o modelo *stepwise mutation* (ver **Glossário de Termos e Pesquisadores**). O R_{ST} capta a fração da variância total do tamanho dos alelos que está entre as populações. O valor de R_{ST} é similar ao valor de θ definido por Weir e Cockerham, uma vez que ambos estão baseados em componentes de variância interpopulacional.

Os índices de fixação podem ser usados para mensurar o padrão de estrutura populacional. A descrição do padrão de fluxo gênico entre um conjunto de subpopulações, contudo, depende de modelos que descrevem os diferentes padrões possíveis.

1.4.2.2. Modelos

O Modelo Continente-Ilhas (Figura 1.23) é o modelo mais simples de fluxo gênico e foi baseado no trabalho de Wright, publicado em 1940 (Wright 1940). Neste modelo assume-se a ligação, por meio do fluxo gênico, de uma população grande, na qual as frequências gênicas mudam pouco ao longo do tempo, com uma (ou mais) população pequena. A cada geração esta população pequena (ilha) recebe imigrantes da população grande (continente). Migração da ilha para o continente também existe, contudo, é desprezível. Neste modelo unidirecional, as frequências alélicas nas populações de ilha tendem a se aproximar daquelas da população continental. A taxa

de migração é semelhante à taxa de mutação em qualidade, mas a taxa de migração é muito maior que a de mutação.

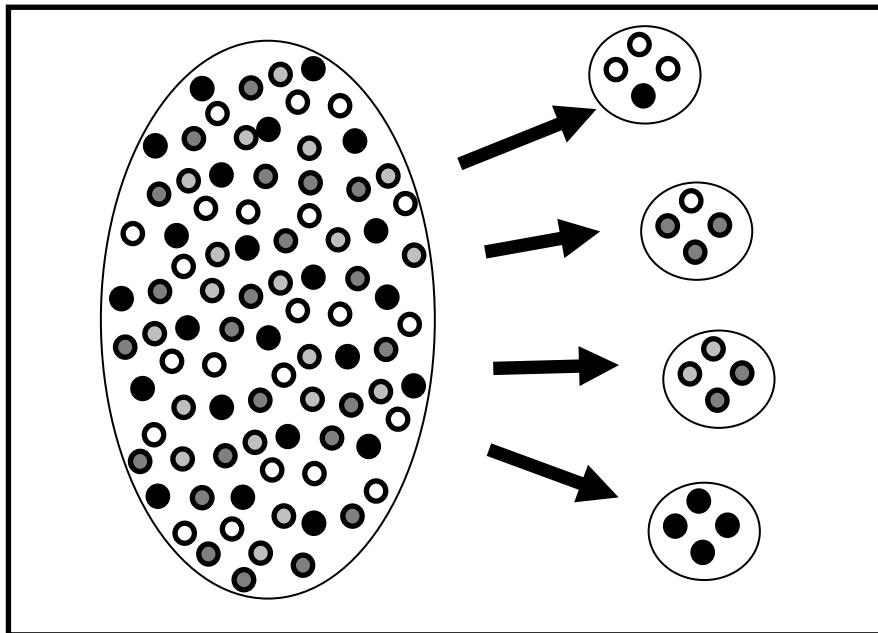


Figura 1.23. Representação do Modelo Continente-Ilhas de estrutura populacional.

No Modelo de Ilhas (Figura 1.24), proposto por Wright, em 1943, existem várias populações, todas elas de tamanho equivalente e a migração ocorre entre elas na mesma taxa. As distâncias entre estas populações não são levadas em conta pelo modelo, o que significa que a migração ocorre ao acaso entre as populações (Wright 1943; 1946).

Neste modelo, a deriva genética pode ser desprezada devido ao fato das populações serem consideradas com grandes tamanhos (populações com mais de 10.000 indivíduos podem ser consideradas panmíticas). Cada subpopulação recebe uma proporção igual de migrantes de todas as demais subpopulações, incluindo ela própria. Essencialmente, neste modelo, cada subpopulação exporta uma proporção de indivíduos (ou gametas) e os migrantes, de todas as subpopulações, formam um conjunto único, cujos membros, então, se dispersam aleatoriamente.

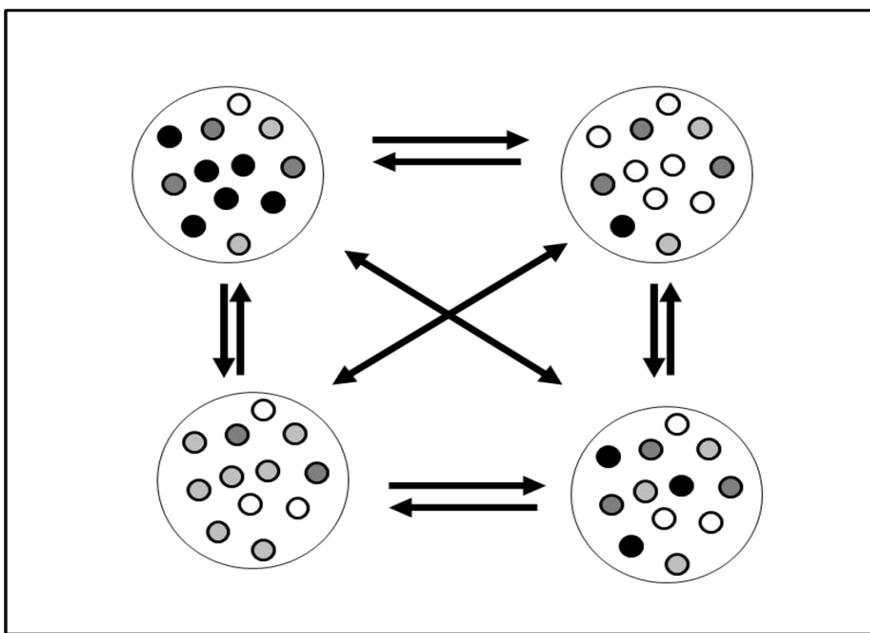


Figura 1.24. Representação do Modelo de Ilhas de estrutura populacional.

Na figura 1.25 está representado o Modelo *Stepping-Stone* (ou, em português, passo-a-passo ou alpondras) que foi proposto por Kimura, em 1953. Porém, tendo sido mais bem explicitado, em 1964, por Kimura e Weiss. Neste modelo, todas as populações, com o mesmo tamanho efetivo e taxa de migração constante, recebem migrantes preferencialmente da população vizinha (Kimura 1953; Kimura & Weiss 1964). Desta forma, o fluxo gênico é mais acentuado entre populações vizinhas do que com populações mais distantes. Dependendo do número de populações vizinhas consideradas, o Modelo *Stepping-Stone* se equipara ao Modelo de Isolamento por Distância, proposto por Wright em trabalhos publicados entre 1940 e 1951. No modelo de Wright, o fluxo gênico ocorre preferencialmente entre indivíduos mais próximos espacialmente, em subpopulações de distribuição contínua.

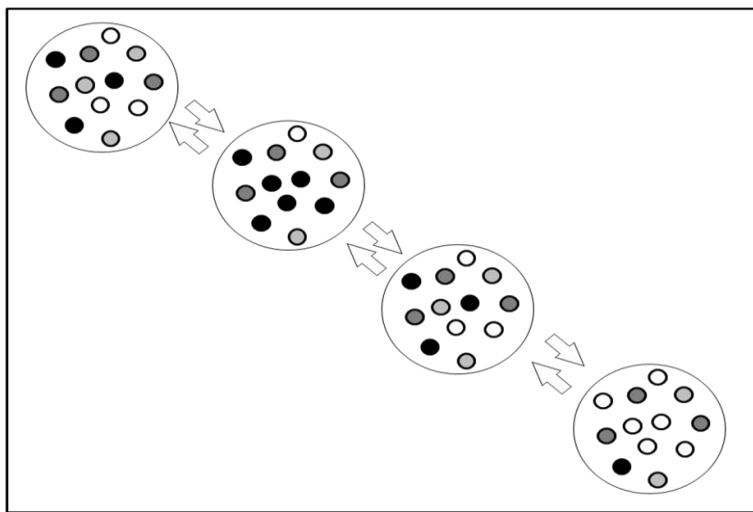


Figura 1.25. Representação do Modelo *Stepping-Stone* de estrutura populacional.

Numa população que se forma pela mistura de subpopulações geneticamente distintas, se observa um fenômeno chamado de efeito Wahlund (Johnson & Black 1984b). Neste caso, como as frequências dos alelos são diferenciadas entre a subpopulações que fornecem migrantes para a nova população, observa-se um grande número de *loci* para os quais a hipótese de equilíbrio de Hardy-Weinberg é violada. Outros efeitos interessantes (e aparentemente paradoxais), de populações formadas pela mistura de subpopulações muito diferenciadas é a existência de altos níveis de variação gênica, contudo, com déficit de heterozigotos. Isto se deve ao fato de que alelos diferentes oriundos de diferentes populações doadoras estão sendo reunidos na população receptora, sem que, contudo, tenham sido formadas todas as combinações possíveis destes alelos em genótipos heterozigotos. Ou seja, o efeito principal da mistura de migrantes de subpopulações diferenciadas é uma diminuição da heterozigosidade média observada em relação aquela esperada se tivessem ocorrido acasalamentos aleatórios nesta população total hipotética.

O efeito Wahlund tanto pode ser tanto de caráter espacial quanto temporal, ou seja, pode ser produzido tanto pela inclusão na mesma amostra de populações de diferentes áreas quanto de diferentes coortes.

1.4.2.3. Relações genéticas

Todos os seres vivos partilham o mesmo material genético que, aqui, tem sido descrito tanto em medidas de fenótipo quanto de genótipo. Estas medidas podem ser usadas para inferir a distância (ou identidade) genética entre indivíduos, populações, espécies ou outras unidades taxonomicas. Ou seja, é possível comparar o percentual de diferença entre genes (ou seus fenótipos, como no caso das aloenzimas) ou sequências de DNA e obter um valor que é a medida de relação genética.

Uma medida de relação genética muito utilizada em Genética Marinha é o índice de Identidade Gênica de Nei, publicado em 1972. Nesta medida, as frequências gênicas são usadas para se calcular a média geométrica das semelhanças padronizadas entre os *loci*. Isto é feito dois a dois. Por exemplo, se se quer saber quais são as identidades gênicas entre três populações (1, 2 e 3), é preciso calcular as identidades entre 1-2, 1-3 e 2-3 (ver Quadro 1.16) (Nei 1972). Quanto mais aparentados geneticamente, ou seja, quanto maior o valor da identidade gênica, menor o tempo de separação entre as duas unidades taxonômicas (OTU, do inglês *Operational Taxonomic Unit*) consideradas. Os valores de identidade gênica podem ser usados, então, para calcular outro índice de relação gênica que é a distância gênica. A distância gênica é calculada através de uma operação matemática simples (ver Quadro 1.17), que é o negativo do cálculo do logarítmico natural do valor de identidade gênica.

Existem outros índices que calculam identidades e distâncias gênicas além do índice proposto por Nei, contudo estes índices foram menos utilizados nos trabalhos de

Genética Marinha. Como exemplos, podem ser citados os índices de Rogers, publicado em 1972 e Thorpe, publicado em 1979 (Rogers 1972; Thorpe 1979; Suzuki & Phan 1990).

Com os valores de identidades (ou distâncias gênicas) dois a dois de todas as unidades taxonômicas trabalhadas é possível construir uma matriz de identidades gênicas (ver Quadro 1.18) entre as OTU'S e, com ela, produzir um dendrograma, que é uma representação gráfica das relações genéticas entre estas unidades taxonômicas. No quadro 1.19, o dendrograma que representa as relações genéticas entre três populações da estrela-do-mar *Coscinasterias tenuispina* é construído passo a passo, usando um dos algoritimos mais simples e singelos que podem ser usados em Genética Marinha, o chamado UPGMA.

UPGMA é uma sigla que significa agrupamento pelas médias aritméticas não ponderadas (em inglês, *Unweighted Pair Group Method with Arithmetic Mean*). O algoritimo UPGMA tem sido largamente utilizado em Genética Marinha, embora existam, na literatura, muitas críticas a ele. Por exemplo, o fato de que esse algorítimo assume que todas as linhagens evoluem a uma taxa constante. Isso significa que os agrupamentos são feitos entre as diferentes OTU's e usando medidas de identidades gênicas calculadas para diferente *loci*, como se tanto os diferentes *taxa* como *loci* evoluíssem a mesma velocidade. Ou seja, como se as taxas de substituição fossem sempre as mesmas independente do *loci* ou do *taxa*. Portanto, a acurácia deste algoritmo depende de que a taxa de substituição se mantenha constante por toda a filogenia.

O pressuposto de regularidade das taxas de substituição entre *taxa* e *loci* é, provalmente, falso para maioria dos casos concretos. Contudo, simulações de computador têm demonstrado que o algoritimo UPGMA é capaz de recuperar adequadamente relações genéticas, especialmente quando utilizado nos níveis

taxonômicos mais inferiores, como relações filogenéticas entre espécies do mesmo gênero. No caso da sua utilização para populações de uma mesma espécie, o algoritimo UPGMA é, ainda, o mais simples, singelo e adequado, não havendo muitas críticas quanto ao seu uso.

Quadro 1.16. Representação matemática da identidade gênica de Nei.

A identidade gênica de Nei pode ser calculada da seguinte forma:

$$I = \sum (f_i \times f_j) / (\sum f_i^2 \times \sum f_j^2)^{0.5}$$

Onde:

f_i- é a freqüência de cada alelo na população i;
f_j- é a freqüência de cada alelo na população j.

Por exemplo, para as três populações de *Coscinasterias tenuispina* descritas no quadro 1.13, para um locus que apresentasse as seguintes frequências gênicas:

População 1:

Locus A: A₁ = 0,70; A₂ = 0,30

População 2:

Locus A: A₁ = 0,80; A₂ = 0,20

População 3:

Locus A: A₁ = 0,50; A₂ = 0,50

As identidade gênicas de Nei serão:

Identidade gênica entre as populações 1 e 2:

$$IG_{1-2} = [(0,70 \times 0,80) + (0,30 \times 0,20)] / \{[(0,70)^2 + (0,30)^2] \times [(0,80)^2 + (0,20)^2]\}^{0,5}$$

$$IG_{1-2} = [0,56 + 0,06] / \{[0,49 + 0,09] \times [0,64 + 0,04]\}^{0,5}$$

$$IG_{1-2} = 0,62 / \{0,58 \times 0,68\}^{0,5}$$

$$IG_{1-2} = 0,62 / \{0,39\}^{0,5}$$

$$IG_{1-2} = 0,62 / 0,628$$

$$IG_{1-2} = 0,987$$

Identidade gênica entre as populações 1 e 3:

$$IG_{1-3} = [(0,70 \times 0,50) + (0,30 \times 0,50)] / \{[(0,70)^2 + (0,30)^2] \times [(0,50)^2 + (0,50)^2]\}^{0,5}$$

$$IG_{1-3} = [0,35 + 0,15] / \{[0,49 + 0,09] \times [0,25 + 0,25]\}^{0,5}$$

$$IG_{1-3} = 0,50 / \{0,58 \times 0,50\}^{0,5}$$

$$IG_{1-3} = 0,50 / \{0,29\}^{0,5}$$

$$IG_{1-3} = 0,50 / 0,538$$

$$IG_{1-3} = 0,928$$

Identidade gênica entre as populações 2 e 3:

$$IG_{2-3} = [(0,80 \times 0,50) + (0,20 \times 0,50)] / \{[(0,80)^2 + (0,20)^2] \times [(0,50)^2 + (0,50)^2]\}^{0,5}$$

$$IG_{2-3} = [0,40 + 0,10] / \{[0,64 + 0,04] \times [0,25 + 0,25]\}^{0,5}$$

$$IG_{2-3} = 0,50 / \{0,68 \times 0,50\}^{0,5}$$

$$IG_{2-3} = 0,50 / \{0,34\}^{0,5}$$

$$IG_{2-3} = 0,50 / 0,583$$

$$IG_{2-3} = 0,857$$

Quadro 1.17. Representação matemática da distância gênica.

A distância gênica pode ser calculada da seguinte forma:

$$D = -\ln IG$$

Onde:

D- é o valor de distância gênica;

ln- logarítmico na base natural;

IG- é o valor de identidade gênica.

Por exemplo, para as três populações de *Coscinasterias tenuispina* que tiveram as seguintes identidades gênicas calculadas no quadro 1.16:

Identidade gênica entre as populações 1 e 2:

$$IG_{1-2} = 0,987$$

Identidade gênica entre as populações 1 e 3:

$$IG_{1-3} = 0,928$$

Identidade gênica entre as populações 2 e 3:

$$IG_{2-3} = 0,857$$

As distâncias gênicas serão:

Distância gênica entre as populações 1 e 2:

$$D_{1-2} = -\ln 0,987$$

$$D_{1-2} = 0,013$$

Distância gênica entre as populações 1 e 3:

$$D_{1-3} = -\ln 0,928$$

$$D_{1-3} = 0,075$$

Distância gênica entre as populações 2 e 3:

$$D_{2-3} = -\ln 0,857$$

$$D_{2-3} = 0,154$$

Quadro 1.18. Representação de uma matriz de identidades e distâncias gênicas.

Por exemplo, para as três populações de *Coscinasterias tenuispina* que tiveram as identidades e distâncias gênicas calculadas nos quadros 1.16 e 1.17:

Identidade e distância gênicas entre as populações 1 e 2:

$$IG_{1-2} = 0,987$$

$$D_{1-2} = 0,013$$

Identidade e distância gênicas entre as populações 1 e 3:

$$IG_{1-3} = 0,928$$

$$D_{1-3} = 0,075$$

Identidade e distância gênicas entre as populações 2 e 3:

$$IG_{2-3} = 0,857$$

$$D_{2-3} = 0,154$$

Teria uma matriz de identidades e distâncias gênicas como a que segue. As identidades estão expressas acima da diagonal e, as distâncias, abaixo da diagonal:

	<i>População 1</i>	<i>População 2</i>	<i>População 3</i>
<i>População 1</i>	***	0,987	0,928
<i>População 2</i>	0,013	***	0,857
<i>População 3</i>	0,075	0,154	***

Quadro 1.19. Representação matemática e gráfica da construção de um dendrograma de UPGMA.

Para o exemplo do quadro 1.16, no qual foram calculadas as identidades gênicas de Nei para três populações de *Coscinasterias tenuispina*:

Identidade gênica entre as populações 1 e 2:

$$IG_{1-2} = 0,987$$

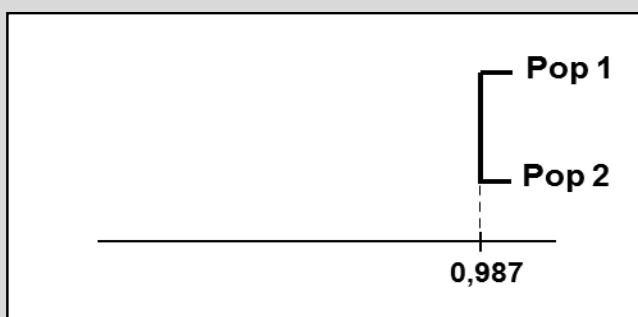
Identidade gênica entre as populações 1 e 3:

$$IG_{1-3} = 0,928$$

Identidade gênica entre as populações 2 e 3:

$$IG_{2-3} = 0,857$$

A maior identidade gênica encontrada é aquela entre a população 1 e a população 2. Com esses dados pode-se começar a montar um dendrograma que refletirá esse padrão de similaridade:



Contudo, agora, com qual identidade gênica a população 3 deve se unir ao grupo formado pela população 1 + população 2? Neste caso, é calculada a média aritmética simples entre as identidades gênicas da população 3 com a população 1 e da população 3 com a população 2. Neste caso:

$$IG_{3-(1-2)} = 0,928$$

$$IG_{3-2} = 0,857$$

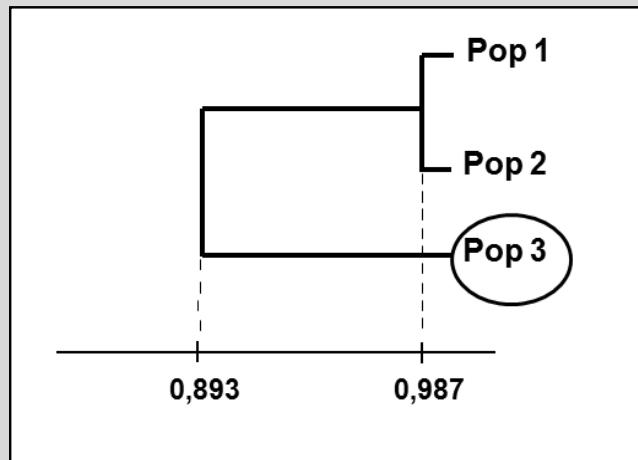
Portanto:

$$IG_{(1-2)-3} = (0,928 + 0,857) / 2$$

$$IG_{(1-2)-3} = 1,785 / 2$$

$$IG_{(1-2)-3} = 0,893$$

Assim, o dendrograma ficará como mostrado abaixo:



Se a regularidade das taxas de substituição entre todos os *taxa* e entre todos *loci* não se sustenta em nenhum caso concreto, ela não é obrigatoriamente falsa para determinados *loci* em determinados *taxa*. Nestes casos, em que a taxa de substituição é regular, ela pode indicar o tempo de divergência entre duas espécies, dois gêneros, duas famílias... Como se fosse o tic-tac de um relógio.

1.4.3. Relógio Molecular

Em 1959, Emile Zuckerkandl (que começou sua carreira como um biólogo marinho) começou um estágio pos-doutoral sob orientação de Linus Pauling, um dos mais renomados bioquímicos do mundo, no *California Institute of Technology*. O trabalho de Zuckerkandl era focado na caracterização de moléculas de hemoglobina de várias espécies utilizando, para tanto, técnicas como a eletroforese e a cromatografia em papel. Os resultados indicaram que o número de diferenças entre os aminoácidos que compunham a hemoglobina das diferentes espécies estudadas seguia um padrão equivalente a distância filogenética entre elas. Ou seja, quanto mais distantes evolutivamente eram as espécies, maior era o número de aminoácidos diferentes entre as suas moléculas de hemoglobina. Contudo, naquele momento esta era apenas uma observação curiosa (Zuckerkandl et al. 1960).

Foi somente em 1962 e, posteriormente, em 1965, que Zuckerkandl e Pauling propuseram a ideia de um relógio molecular (Zuckerkandl & Pauling 1962; 1965). Comparando o número de diferenças de aminoácidos nas cadeias α e β de hemoglobina e usando as evidências paleontológicas do tempo de divergência entre cavalos e humanos, eles foram capazes de inferir, com base no número de substituições de aminoácidos entre as diferentes moléculas, o tempo de divergência entre várias espécies de mamíferos desde o último ancestral entre elas. Tudo isso, assumindo que as

mudanças de aminoácidos entre as moléculas de proteínas iam sendo acumuladas de maneira constante ao longo do tempo. Estava, então, proposta mais uma ideia controversa no seio da mais controversa de todas as teorias, que é a teoria evolutiva.

Não por acaso, os dois trabalhos no quais Zuckerkandl e Pauling desenvolveram a ideia do relógio molecular, não foram artigos que tivessem sofrido a análise estrita do duplo cego a que são, geralmente, submetidos os artigos científicos. Em ambos os casos, foram capítulos de livros que os autores tinham sido convidados a escrever e que, desta forma, pelas suas próprias palavras, eles puderam usar a oportunidade para dizer algo tão escandaloso como era a ideia de um relógio molecular (*"we took the opportunity to say something outrageous"*).

Basicamente, o relógio molecular é uma hipótese que propõe que os genes evoluem a uma taxa constante ao longo do tempo. Desta forma, esta hipótese traz a possibilidade de inferir o tempo de divergência entre duas linhagens a partir do número de substituições entre as suas sequências de proteínas e/ou DNA (Morgan 1998; Thomas et al. 2006; Thomas et al. 2010; Calcagnotto 2012; Ho & Duchêne 2014). Obviamente, as aplicações do relógio molecular se tornaram, rapidamente, comuns em estudos de Genética Marinha, especialmente nos casos em que eventos como a subida do Istmo do Panamá ou o surgimento do Arquipélago do Havaí possibilitava saber o tempo de separação entre os *taxa*.

1.4.3.1. Taxa de substituição

A taxa de substituição em nível molecular é dada pela quantidade de divergência entre sequências ocorrida por unidade de tempo. Assim, se duas sequências forem comparadas e se for sabido há quanto tempos elas divergiram de uma sequência

ancestral comum, é possível estimar a taxa de substituição para estas sequências (Figura 1.26).

As unidades são expressas como substituições por sítio de aminoácido (ou substituições por sítio de nucleotídeo) por ano. Contudo, em longos períodos do tempo evolutivo, os aminoácidos presentes em um sítio em particular podem ter sido substituídos diversas vezes. Primeiramente por um aminoácido qualquer, depois por outro, depois outro e, ainda, outro. Até que, em algum momento, este processo determine a volta do aminoácido originalmente presente naquele sítio. Ao serem comparadas duas sequências, apenas os sítios que são diferentes podem ser identificados. Portanto, para longos tempos evolutivos, é possível que alguns dos sítios que são idênticos no presente possam ter sido diferentes no passado. Do mesmo modo, sítios que são diferentes no presente, podem ter sofrido mais do que uma substituição. Para intervalos de tempo relativamente curtos, durante os quais múltiplas substituições são, ainda, incomuns, o problema é pequeno. Contudo, para longos intervalos de tempo, este problema se torna complicado.

O problema das substituições múltiplas é mais complicado, ainda, no caso de nucleotídeos. Para aminoácidos, a probabilidade de que uma substituição ao acaso faça um sítio retornar ao seu estado original é de 1 para 20 (que é o número de aminoácidos). No caso de nucleotídeos, a chance de substituições múltiplas é de 1 para 4 (que é o número de bases, A, T, G e C). Mesmo estas estimativas são simplificações, uma vez que, em ambos os casos, se esta assumindo o pressuposto de que a chance de mudança de um aminoácido para o outro (ou de um nucleotídeo para o outro) é igual entre todos os casos.

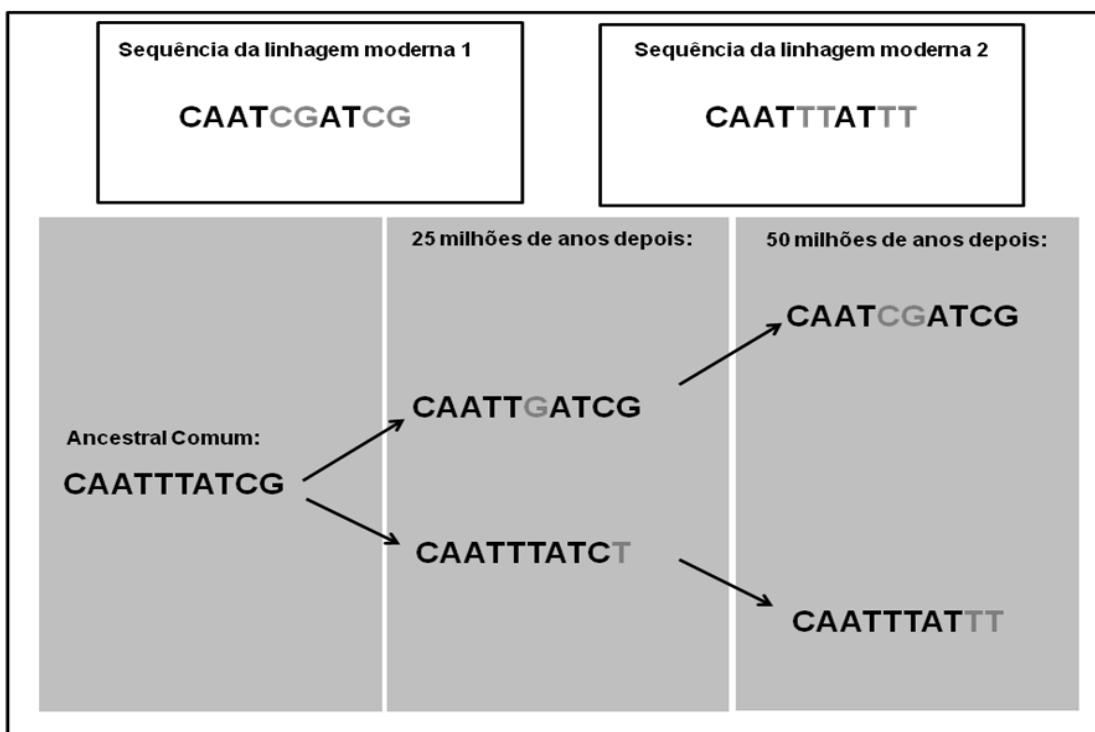


Figura 1.26. Há 50 milhões de anos atrás, as duas linhagens modernas tiveram um ancestral comum. Após 25 milhões de anos, as duas linhagens divergiram, cada uma delas tendo uma mudança de base. De modo que ficaram diferentes por duas bases. Após outros 25 milhões de anos, estas duas sequências acumularam mais duas bases de diferença (um nucleotídeo mudado para cada uma delas). Assim, após o total de 50 milhões de anos as duas linhagens divergiram em quatro bases (duas em cada linhagem). Segundo a hipótese do relógio molecular este processo de divergência é constante e regular e, no exemplo ilustrado, o seu tic-tac é de uma mudança de base a cada 25 milhões de anos, portanto, depois de 50 milhões de anos as duas sequências acumularam quatro diferenças de base entre si e, cada uma delas, duas diferenças em relação ao ancestral comum. O que significam 100 milhões de anos de divergência entre as sequências.

O pressuposto de que as substituições ocorrem com igual probabilidade entre todos os aminoácidos (ou nucleotídeos) é desmentido por uma grande quantidade de dados. Além do fato de que proteínas reais violam esse pressuposto, não era esperado que ele fosse verdadeiro, uma vez que algumas substituições de aminoácido requerem apenas uma única substituição de nucleotídeo, enquanto outras requerem duas ou mesmo três substituições. Modelos mais sofisticados para a evolução de aminoácidos levam em conta essas diferenças ao dar pesos para cada substituição de aminoácido de

acordo com a taxa observada dessas substituições em um grande número de proteínas. No caso das sequências de nucleotídeos as correções são mais importantes ainda, visto que é esperado que um terço das substituições ao acaso faça com que dois sítios diferentes se tornem iguais.

Diversos modelos de substituição de nucleotídeos têm sido propostos e estudados, os quais diferem, basicamente, nos pressupostos quanto à taxa de substituição entre pares de nucleotídeos. O modelo mais simples é o de Jukes-Cantor, proposto em 1969, no qual as substituições ocorrem a uma taxa constante e cada nucleotídeo tem uma chance igual de mutar para qualquer outro. Um modelo mais complexo do que este foi proposto, em 1980, por Kimura. No seu modelo Kimura considera o fato de que comparações de sequências têm demonstrado que a frequência de transições (ou seja, mudanças de uma purina para outra purina-G para A; ou de uma pirimidina para outra pirimidina-C para T) é maior do que a de transversões (qualquer outra mudança possível). Neste caso, dois parâmetros para a taxa de mutação são considerados. As transições ocorrem com uma frequência α , enquanto as transversões ocorrem com taxa β . Outros modelos, ainda, foram propostos adicionando diferentes parâmetros às diferentes possibilidades de substituição. Um dos mais populares é o que permite quatro taxas de substituição diferentes, embora, existam modelos ajustados para, até, 12 taxas de substituição diferentes. Contudo, as vantagens de modelos tão complexos são discutíveis, especialmente nos casos em que modelos mais simples oferecem resultados não muito diferentes daqueles obtidos pelos modelos de muitos parâmetros.

A grande lição disto tudo é que, a despeito do relógio molecular ser amplamente utilizados em estudos evolutivos no geral, e em Genética Marinha em específico, ele é,

de fato, uma grande simplificação. Os dados têm demonstrado que existe uma heterogeneidade muito grande de variáveis relacionadas às taxas de substituição.

1.4.3.2. Taxa das sequências

As taxas de substituição não são universalmente constantes. Por exemplo, as substituições de nucleotídeo ocorrem a uma taxa maior (1 em 4) do que as de aminoácidos (1 em 20), como já foi visto. No caso dos nucleotídeos, as substituições ocorrem em diferentes taxas dependendo da região do DNA onde ocorrem, dependendo do tipo de genoma e, também, se ela é sinônima ou não.

Devido à redundância do código genético (ver Quadro 1.20), dos vinte aminoácidos, oito possuem qualquer nucleotídeo na sua terceira base e são chamados de sítios quatro vezes degenerados, porque a ocorrência de qualquer um dos quatro nucleotídeos é indiferente. Sete nucleotídeos terminam em qualquer pirimidina (o que significa T ou C) e cinco nucleotídeos terminam em qualquer purina (o que significa A ou G) e são chamados de sítios duas vezes degenerados. Em virtude disto (as chamadas degenerações), os nucleotídeos em um gene podem mudar sem afetar a sequência de aminoácidos da proteína que codificam. Essas mudanças são chamadas de substituições de nucleotídeo sinônimas ou silenciosas. Substituições de nucleotídeos que trocam aminoácidos são substituições não sinônimas. As taxas sinônimas são, em geral, muito maiores do que as taxas de substituição em sítios não sinônimos. A depressão na taxa de substituição não sinônica é interpretada como sendo causada por seleção natural, a qual elimina aquelas mudanças que são deletérias (Hipótese Neutralista, ver *1.2.1.3. Controvérsias evolutivas*).

Quadro 1.20. Código genético com a apresentação da redundância de várias das trincas (conjunto de três bases).

Diz-se que o código genético é degenerado (ou redundante) porque existem vários codons que codificam para o mesmo aminoácido. Por exemplo, os codons UCU, UCC, UCA e UCG codificam todos eles o aminoácido Serina (Ser). Neste caso, pode-se dizer que Serina é um aminoácido quatro vezes degenerado. Todos os códons quatro vezes degenerados estão marcados em branco.

		2 ^a base			
		U	C	A	G
1 ^a base	U	UUU - <u>Fenilalanina</u> UUC - <u>Fenilalanina</u> UUA - <u>Leucina</u> UUG - <u>Leucina</u>	UCU - <u>Serina</u> UCC - <u>Serina</u> UCA - <u>Serina</u> UCG - <u>Serina</u>	UAU - <u>Tirosina</u> UAC - <u>Tirosina</u> UAA - Códon de parada UAG - Códon de parada	UGU - <u>Cisteína</u> UGC - <u>Cisteína</u> UGA - Códon de parada UGG - <u>Triptofano</u>
	C	CUU - <u>Leucina</u> CUC - Leucina CUA - <u>Leucina</u> CUG - <u>Leucina</u>	CCU - <u>Prolina</u> CCC - <u>Prolina</u> CCA - <u>Prolina</u> CCG - <u>Prolina</u>	CAU - <u>Histidina</u> CAC - <u>Histidina</u> CAA - <u>Glutamina</u> CAG - <u>Glutamina</u>	CGU - <u>Arginina</u> CGC - <u>Arginina</u> CGA - <u>Arginina</u> CGG - <u>Arginina</u>
	A	AUU - <u>Isoleucina</u> AUC - Isoleucina AUA - <u>Isoleucina</u> AUG - <u>Metionina</u>	ACU - <u>Treonina</u> ACC - <u>Treonina</u> ACA - <u>Treonina</u> ACG - <u>Treonina</u>	AAU - <u>Asparagina</u> AAC - <u>Asparagina</u> AAA - <u>Lisina</u> AAG - <u>Lisina</u>	AGU - <u>Serina</u> AGC - <u>Serina</u> AGA - <u>Arginina</u> AGG - <u>Arginina</u>
	G	GUU - <u>Valina</u> GUC - <u>Valina</u> GUA - <u>Valina</u> GUG - <u>Valina</u>	GCU - <u>Alanina</u> GCC - <u>Alanina</u> GCA - <u>Alanina</u> GCG - <u>Alanina</u>	GAU - <u>Ácido aspártico</u> GAC - <u>Ácido aspártico</u> GAA - <u>Ácido glutâmico</u> GAG - <u>Ácido glutâmico</u>	GGU - <u>Glicina</u> GGC - <u>Glicina</u> GGA - <u>Glicina</u> GGG - <u>Glicina</u>

O padrão de transmissão hereditária afeta, também, a taxa de substituição. Os genomas de organelas são transmitidos de uma maneira diferente dos genes nucleares, e, portanto, sofrem substituições com uma dinâmica diferente. Mesmo considerando apenas um tipo de DNA, por exemplo, o mtDNA, é possível observar uma ampla variação nas taxas de substituição ao longo do seu pequeno genoma. Em animais, a taxa de substituição no mtDNA é, em geral, muito mais alta do que aquela exibida no nDNA. Em plantas, comparações entre as taxas de substituição de nucleotídeo entre o DNA nuclear, de cloroplasto e mitocondrial revelam, também, grandes diferenças, sendo que o mtDNA exibe menos do que um terço da taxa de substituição do cpDNA, o qual por sua vez apresenta cerca de metade da taxa de substituição do nDNA.

Em princípio, o relógio molecular deveria aplicar-se tanto às regiões intergênicas (não codificantes) do genoma quanto aos genes em si. Entretanto, para maioria esmagadora dos casos, a taxa de substituição é muito maior para os introns (regiões não codificantes) do que para os exons (regiões codificantes). A explicação para tal fato, mais uma vez, está na teoria neutralista que afirma que as regiões codificantes estão sofrendo a ação da seleção natural e, portanto, sofrem restrições às mudanças, que só podem ocorrer no caso de não causarem danos aos genes. Nas regiões não codificantes, como não existem restrições impostas pela seleção natural (uma vez que elas não codificam nada!), a mudança é livre.

Como resultado da grande variação entre as taxas de substituição, foi proposta a idéia de relógios moleculares locais. Ou seja, relógios específicos para determinados genomas e suas regiões. Contudo, ainda resta o problema de saber se, assumidas estas restrições, a taxa de substituição é constante entre os diferentes grupos taxonômicos.

1.4.3.3. Taxa dos taxa

A existência de um relógio molecular universal foi colocada em dúvida devido ao fato de que as taxas de substituição variam amplamente entre os diferentes *taxa*. Ou seja, existe um grande número de evidências que indicam que as taxas de substituição podem variar consideravelmente entre as espécies para uma grande gama de *taxa*. O relógio molecular não bate a um ritmo constante, mesmo entre linhagens estreitamente relacionadas!

Diante deste problema, métodos para estimar a variação da taxa de substituição entre *taxa* foram desenvolvidos. Contudo, em todos os casos, os pressupostos destes métodos são, inevitavelmente, *ad hoc* e sem base empírica. Ou seja, a confiança nesses

métodos depende de se compreender o que determina a heterogeneidade das taxas de substituição para os diferentes *taxa*, o que não se sabe, portanto.... É preciso dizer mais?

Muitos métodos de datação molecular assumem que a variação na taxa de substituição é, essencialmente, estocástica. Por outro lado, estima-se que as taxas variam sistematicamente entre as linhagens, ou seja, seriam determinadas pelas características das espécies. Existe uma gama de correlações entre taxa de substituição e história de vida das espécies, tanto intrínsecas (por exemplo, taxa metabólica) quanto extrínsecas (por exemplo, fatores ambientais como temperatura), e, mesmo, processos evolutivos como o tempo de especiação no *taxa*. Contudo, a melhor caracterização de uma associação sistemática entre as características das espécies e taxa de substituição é o efeito do tamanho corporal, observada para vertebrados. Espécies com corpo menor tendem a ter taxas mais rápidas da evolução molecular do que as espécies próximas com corpo maior. Para invertebrados, no entanto, não existem boas caracterizações. O efeito do tamanho do corpo dos vertebrados é um caso especial, que não pode ser simplesmente extrapolados para todo os *taxa*.

Com todas estas limitações (heterogeneidade das taxas de substituição para diferentes genomas e suas regiões e entre taxa), é evidente que se deve tomar muito cuidado na escolha do relógio a ser usado, sob o risco de estimativas altamente enganosas de tempos de divergência. Assim, um relógio para cada um e para cada lugar. Mais que isso, é preciso acertar o relógio, ou seja, escolher um ponto adequado de calibração.

1.4.3.4. Pontos de calibração

Calibrar é estabelecer, sob condições específicas, a relação entre os valores de um sistema de medição e os valores de um material de referência. No caso do relógio molecular, acertar a taxa de substituição ao tempo.

Um dos métodos mais confiáveis de calibração é a utilização de evidências fósseis que indiquem o tempo de divergência entre duas ou mais linhagens incluídas na topologia que se pretende usar para estimar os tempos de divergência. No entanto, mesmo quando há disponibilidade de fósseis que definem os tempos de separação entre *taxa*, fontes de erros na inferência filogenética precisam ser consideradas. Alternativamente, pode-se lançar mão de eventos geológicos associados à diversificação de espécies de um dos grupos na topologia. Por exemplo, a origem de ilhas oceânicas cuja data de emergência é conhecida. Porém, da mesma forma que com o registro fóssil, existem erros de inferência associados a estas estimativas como, por exemplo, a precisão do intervalo de tempo para a formação da ilha, o tempo de colonização das populações após a formação de cada ilha e a possibilidade de migração entre as ilhas.

Por causa dos erros associados a qualquer calibração (o que se torna mais dramático com o aumento da distância entre os nós de separação entre os *taxa*), o ideal é usar uma abordagem múltipla, na qual vários pontos de calibração são usados simultaneamente. Ou seja, um relógio para cada um, um relógio para cada lugar e, fundamentalmente, não esquecer que conferir cada um contra todos os outros. Afinal, pior do que tantos relógios é relógio algum.

1.4.4. Coalescência

Tudo que foi dito até aqui se baseou nas duas novidades da teoria evolutiva darwiniana: a perspectiva materialista da variação e, a partir dela, a possibilidade de entender o processo de especiação como um processo de transformação de variação intrapopulacional em variação interpopulacional. Mais que isso, todas as medidas e estimativas discutidas até agora foram baseadas nos desenvolvimentos matemáticos (especialmente de Wright e Fisher) da teoria sintética da evolução. Contudo, um dos corolários mais importantes da teoria evolutiva darwiniana é a ancestralidade comum. É neste corolário que se baseia a teoria da coalescência.

Coalescência é o processo pelo qual, olhando para trás no tempo, as genealogias de dois alelos se unem em um ancestral comum (Fernandes 2012). Assim, ao invés de se olhar o processo de diversificação, olha-se retrospectivamente e vê-se a diversidade de alelos do presente coalescer ao alelo original. Este alelo original é o que se chama na teoria da coalescência de ancestral comum mais recente (ou MRCA, do inglês *most recent common ancestor*). Desta forma, em uma amostra de “n” alelos, por exemplo, o primeiro evento de coalescência une as “n” genealogias atuais em “n-1” genealogias ancestrais, o segundo evento de coalescência une em “n-2” genealogias e assim por diante, até que sobre um único ancestral comum para todo o conjunto de alelos amostrados (Figura 1.27).

A cada geração, existe um número de alelos na população, os quais podem ser replicados e estar presentes na próxima geração, em alguns casos, porém, um alelo não deixa descendente e é perdido da população. Com a repetição desse processo ao longo do tempo, um dos alelos originais acabará se tornando fixado na população. Na ausência de mutação, a população acabaria, portanto, com o mesmo alelo fixado. Entretanto,

como pode ocorrer mutação durante esse processo, os alelos observados no presente não serão todos idênticos na sua sequência de nucleotídeos, embora sejam todos descendentes de um único alelo ancestral comum.

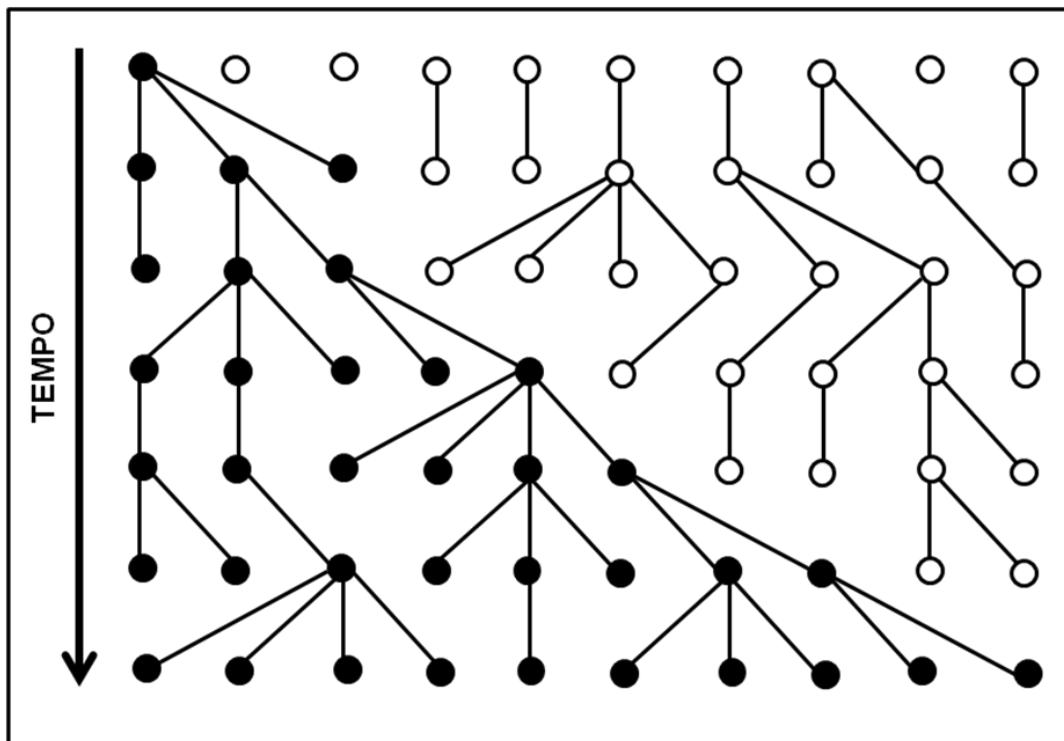


Figura 1.27. Esquema representativo das rotas de ancestralidade de um conjunto de alelos. À medida que as gerações avançam no tempo (de cima para baixo), muitos alelos não deixam descendentes e, assim, se extinguem (círculos vazios). No processo ao contrário (de baixo para cima), a amostra de alelos observada na geração atual (círculos cheios) sofre uma série de eventos de coalescência nos quais os “n” alelos presentes na geração atual têm apenas “n – 1” alelos ancestrais até o ancestral comum mais recente.

Em princípio a teoria da coalescência parece uma troca de seis por meia dúzia. Ou seja, troca-se uma interpretação do processo como sendo diversificador (variação intrapopulacional gerando variação interpopulacional), para uma interpretação do processo como sendo coalescente (ancestralidade comum). Contudo, o poder da teoria da coalescência está no fato de que para determinadas análises, ela economiza muito tempo computacional. Por exemplo, na figura 1.27 se se quisesse seguir a evolução dos alelos do passado para o presente (processo diversificador), seria preciso seguir o

destino de todos os setenta alelos representados, mesmo aqueles 34 alelos que se extinguiram pelo meio do caminho (círculos vazios). Por outro lado, usando a teoria da coalescência, apenas o destino dos 36 alelos que coalescem para o ancestral comum mais recente (círculos cheios) seriam relevantes para análise. Isto significa 50% de esforço a menos, o que não é pouco!

Estudos de coalescência procuram entender as relações genealógicas, ou seja, a história evolutiva de determinados genes dentro de uma espécie particular. Além das forças evolutivas (seleção natural, deriva genética, mutação e migração), muitos processos históricos (colonização, extinção, dispersão, expansão populacional) também influenciam a distribuição geográfica de linhagens genealógicas dentro de uma mesma espécie ou mesmo entre espécies próximas. O estudo das relações entre genealogias, fatores históricos e geografia é o objeto de estudo da filogeografia.

1.4.5. Filogeografia

Filogeografia se define como o estudo dos processos históricos que influenciam a distribuição geográfica contemporânea de indivíduos, sejam eles da mesma espécie ou de espécies próximas. Assim, considera-se tanto a distribuição geográfica dos indivíduos quanto os padrões genealógicos de um ou alguns genes (geralmente mtDNA). O foco na distribuição geográfica de uma espécie, ou seja, no seu padrão biogeográfico, é o que confere a filogeografia a sua especificidade, distinguindo-a da genética de populações e dos estudos filogenéticos estritos. Os eventos históricos que interessam diretamente a filogeografia são a expansão populacional, efeitos de gargalo, vicariância e migração, uma vez que são eles que têm efeito direto sobre os padrões atuais de distribuição geográfica dos indivíduos das espécies (Martins & Domingues 2011; Cunha & Solé-Cava 2012).

Os estudos nesta disciplina foram inaugurados por John Avise, em 1979, quando, trabalhando com a genética de populações de roedores da espécie *Geomys pinetis*, resolveu, juntamente com seus colaboradores, usar os indivíduos, ao invés das populações, como OTU's (Avise et al. 1979). Mais que isso, construíram redes de haplótipos para os padrões de RFLP obtidos para mtDNA, as quais foram sobrepostas aos locais de amostragem. O resultado desta estratégia foi que eles conseguiram evidenciar diferenças entre as populações coletadas ao leste e a oeste dos Estados Unidos que apresentavam uma concordância com o padrão de glaciações do Pleistoceno. Bingo! Surgia aí uma análise que incorporava dados de relações genéticas e posicionamento geográfico, uma filogeografia.

Em um artigo publicado em 1987, Avise e outros pesquisadores lançariam as bases teóricas e metodológicas desta nova disciplina. O título do artigo já dizia tudo: *Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics* (Filogeografia intraespecífica: A ponte que o mtDNA faz entre a genética de populações e a sistemática) (Avise et al. 1987). Ou seja, a biogeografia que já existia e estudava como as condições geológicas, climáticas e ecológicas influenciavam a distribuição atual das espécies recebia um reforço da genética de populações, dos marcadores moleculares e da teoria da coalescência. Aos poucos, somaram-se a isso, ainda, as análises cladísticas e as teorias de tectônica de placas e vicariância. A filogeografia chegava a sua maioridade. Hoje, amadurece como uma disciplina prestigiosa que tem contribuído imensamente para Genética Marinha.

Um exemplo clássico da utilização da filogeografia em Genética Marinha diz respeito aos estudos com ouriços-do-mar na região oeste do Oceano Indo-Pacífico. Esta região é, reconhecidamente, um centro de elevada diversidade marinha que declina acentuadamente em direção aos extremos do arquipélago indonésio. Este fato fez com

que pesquisadores levantassem a hipótese de que esta região era um centro de origem, acumulação e dispersão de biodiversidade. Estudos de Filogeografia, realizados na década de 1990, permitiram evidenciar que as espécies do gênero *Echinometra* são de origem recente, tendo divergido dentro do Pleistoceno e a hipótese de que o oeste do Oceano Indo-Pacífico é um centro de origem, acumulação e dispersão de biodiversidade não foi confirmada, sugerindo que a formação de espécies pode ocorrer ao longo de todo Oceano Indo-Pacífico (Palumbi 1996).

Como já deve ter ficado evidente, o marcador molecular preferido em estudos de filogeografia é o mtDNA. Isto se deve ao fato de que esta molécula de DNA possui herança uniparental e não apresenta recombinação. Contudo, análises baseadas em árvores de mtDNA representam apenas um *locus* e, portanto, são muito limitadas, podendo gerar conclusões equivocadas quanto a história evolutiva dos *taxa*. Desta forma, os estudos de filogeografia têm tentado combinar marcadores mitocondriais e nucleares nos seus testes de hipóteses filogeográficas.

Alguns padrões filogeográficos podem ser categorizados de acordo com o grau de separação genética e geográfica das linhagens. De uma maneira geral, são reconhecidos quatro tipos de estruturas filogeográficas que formam um contínuo no tempo, desde a separação mais antiga até a mais recente. A primeira categoria é chamada de categoria I, na qual são agrupadas as separações mais antigas, com grandes divergências genéticas, profundamente estruturadas. Na categoria II estão agrupadas as separações um pouco menos antigas, com grandes divergências genéticas e estruturadas. A categoria III não apresenta grandes divergências genéticas e nem estruturação. Por fim, a categoria IV apresenta as menores divergências genéticas. Para as categorias I e II, o mtDNA é uma ótima alternativa, uma vez que se trata de separações mais antigas, com separação completa das linhagens no período das ramificações e por isso,

geralmente, as árvores de genes e populações apresentam o mesmo padrão (Figura 1.28). Nas categorias III e IV é preciso utilizar mais *loci* do que é oferecido pelo mtDNA, uma vez que as divergências genéticas são menores. Nestes casos (populações pouco estruturadas geograficamente), as inferências sobre história populacional recente se beneficiam bastante da utilização de análises filogenéticas de coalescência, com definição *a posteriori* das populações. Geralmente, nas categorias III e IV, as árvores de genes e populações não apresentam o mesmo padrão (Figura 1.28).

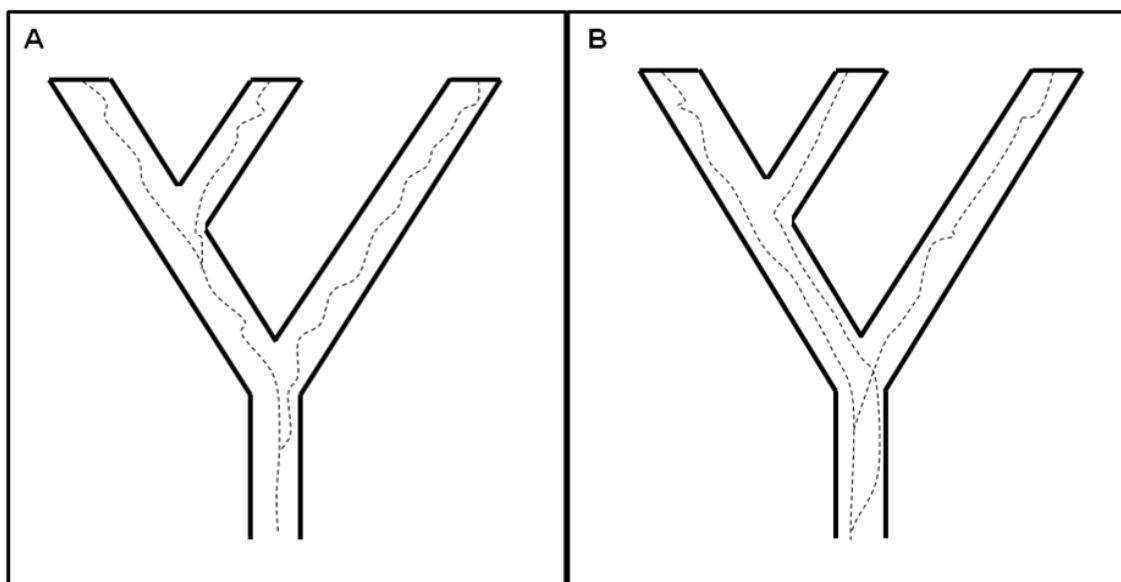


Figura 1.28. Árvores de espécies (contorno em traço contínuo) e árvores de genes (linha pontilhada). A- A topologia da árvore de espécies e da árvore de genes são iguais; B- A topologia da árvore de espécies e da árvore de genes são diferentes.

Os primeiros trabalhos filogeográficos possuíam uma natureza de narração de fatos e não apresentavam rigor estatístico (não apresentava testes de hipóteses alternativas). Mais recentemente, os trabalhos tem se tornado mais sofisticados, tentando determinar a validade de processos e explicar a concordância entre distância geográfica e a proximidade genética.

1.5. COMPUTACIONAL

Dependendo da massa de dados obtidos empiricamente, a interpretação conceitual pode ser muito beneficiada pelo auxílio de computadores que, de fato, não são apenas máquinas, mas um sistema, chamado computacional. Este sistema computacional é o produto da interação entre dispositivos eletrônicos (a parte física ou *hardware*, em inglês), programas (a parte lógica ou *software*, em inglês) e os pesquisadores (programação ou uso, *peopleware*, em inglês). Ou seja, não há dúvida que, neste sistema, são os pesquisadores que impõem a intenção sobre uma determinada massa de dados de forma a produzir as informações e resultados de interesse (Althusser 1972; 1980).

Na atualidade, esse sistema de manipulação de dados e informações está sofrendo uma revolução no tempo e eficiência de resposta e produzindo riqueza e, ao mesmo tempo, se desenvolvendo em resposta as demandas do sistema de produção capitalista. Ou seja, uma revolução que demonstra a relação dialética entre desenvolvimento dos meios e relações de produção na sociedade (Marx & Engels 1848; Engels 1883; Levins & Lewontin 1985). No caso do sistema computacional isto vem acontecendo desde a era dos dispositivos mecânicos (3000 a.C. até 1880, começando com o desenvolvimento, pelos chineses, do ábaco), passando pela era dos dispositivos eletromecânicos (1880-1930, com o desenvolvimento das máquinas de somar acionadas por motores elétricos), até chegar a esta era atual ou dos computadores (equipamentos eletrônicos de processamento de dados) de que trata especificamente o computacional referido aqui. Nesta revolução, a Genética Marinha menos contribui e mais usufrui.

1.5.1. Revolução Computacional

A revolução computacional inclui o avanço tecnológico, o desenvolvimento de métodos numéricos para o processamento de grandes conjuntos de dados e a utilização

da internet como meio de comunicação e como um repositório de dados. O avanço tecnológico é visível no desenvolvimento técnico dos computadores pessoais (*Personal Computer-PC*, em inglês). Os PC's atuais ultrapassam a capacidade dos processadores que eram exclusivos de grandes laboratórios há 30 anos.

Durante o início dos estudos empíricos de genética de populações (e, portanto, da Genética Marinha), na década de 1970, as análises dos dados eram feitas em programas privados desenvolvidos por pesquisadores individuais que, posteriormente, circulavam de mão em mão entre os demais. Neste período, o conhecimento de linguagens de programação (como o *Fortran*, por exemplo) e o acesso a equipamentos especializados eram necessários para criar, programar e/ou usar os procedimentos estatísticos necessários. Hoje, uma grande diversidade de programas está disponível para uso imediato (muitos dos quais disponíveis gratuitamente na “rede”) e são amplamente utilizados pelos pesquisadores (Tabela 1.2).

Da mesma forma, o desenvolvimento da capacidade computacional e sua acessibilidade, tornaram possível aos pesquisadores a utilização de métodos que permitem a manipulação de conjuntos gigantescos de dados, viabilizando inferências estatísticas em uma escala inimaginável há pouco tempo atrás. Este é o caso, por exemplo, dos testes estatísticos com base em probabilidades e métodos bayesianos que exigem milhões de iterações.

Com o desenvolvimento concomitante das possibilidades computacionais, numéricas e empíricas (chegada ao campo pelas plataformas NGS, ver 4.3.6.5. *Sequenciamento*), ocorreu uma explosão na quantidade de informações produzidas que não poderia ser utilizada sem o acesso direto e fácil proporcionado pelos bancos de dados organizados e disponibilizados na *World Wide Web* (termo em inglês que, em português, se traduz como algo parecido como "Teia mundial de interligação"). Para

Genética Marinha os mais importantes deles são, provavelmente, aqueles mantidos pelo centro nacional de informação tecnológica, o NCBI (*National Center for Biotechnology Information*, em inglês).

O NCBI foi criado, em 1988, pelo governo norte americano, como fonte de referência de informações em Biologia Molecular. O NCBI possui bancos de dados públicos e faz intercâmbio com outros centros internacionais de pesquisa, como os bancos europeu (EMBnet- *The European Molecular Biology Network*) e japonês (DDBJ- *DNA Database of Japan*). Entre os bancos de dados disponíveis no NCBI podem ser encontrados àqueles referentes à sequências de genes (incluindo introns e exons), complementares (oriundas da transcrição reversa de RNA), proteínas e outros. Além disso, muitos aplicativos (por exemplo, aqueles referentes à comparação e busca de sequências) estão disponíveis nesta plataforma. É fácil imaginar como esta nova realidade (real e virtual) beneficiou o trabalho de pesquisa em biologia molecular, genética de populações e, portanto, da Genética Marinha.

Tabela 1.2. Softwares, sua função, sítios onde estão disponíveis e referência original de apresentação do programa.

Software	Breve descrição	Sítios	Referência
<i>Biosys</i>	Calcula índices básicos de diversidade genética e estatística F; teste exato de Equilíbrio de Hardy-Weinberg.	http://fortinbras.us/biosys2/	Swofford, D.L. & Selander, R.B. 1981. BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. <i>Journal of Heredity</i> 72:281-283.
<i>GENEPOP</i>	Calcula índices básicos de diversidade genética e estatística F; teste exato de Equilíbrio de Hardy-Weinberg, de desequilíbrio de ligação; teste de Mantel.	http://ftp.cefe.cnrs.fr/PC/MSDOS/GENEPOP	Raymond, M. & Rousset, F. 1995. GENEPOP (version 3.3): population genetics software for exact tests and ecumenicism. <i>Journal of Heredity</i> 86:248-249.
<i>FSTAT</i>	Calcula índices básicos de diversidade genética, riqueza de alelos e estatística F; teste exato de Equilíbrio de Hardy-Weinberg, de desequilíbrio de ligação; análise de regressão múltipla; teste de Mantel.	http://www2.unil.ch/popgen/softwares/fstat.htm	Goudet, J. 1995. FSTAT (Version 1.2): A Computer program to calculate F-statistics. <i>The Journal of Heredity</i> 86(6):485-486.
<i>Arlequin</i>	Calcula índices de diversidade gênica, estatística F, distâncias genéticas entre populações; teste exato de Equilíbrio de Hardy-Weinberg, de desequilíbrio de ligação e diferenciação da população; testes seletivos e de neutralidade dentro das populações; teste de Mantel; estima parâmetros demográficos de distribuição.	http://cmpg.unibe.ch/software/arlequin3/	Schneider S, Roessli D and Excoffier L (2000) <i>Arlequin: a software for population genetics data analysis. User manual ver 2.000</i> . Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva.
<i>Structure</i>	Usa inferência Bayesiana para, a partir dos genótipos multilocios dos indivíduos, estimar o número de populações que melhor explica os dados, ao mesmo tempo em que os indivíduos são atribuídos às populações.	http://pritch.bsd.uchicago.edu/software/structure2_1.html	Pritchard, J.K.; Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. <i>Genetics</i> 155(2):945-959.
<i>MrBayes</i>	Usa inferências bayesianas para escolha de modelos filogenéticos baseados em dados de sequências de DNA.	http://mrbayes.csit.fsu.edu/	Ronquist, F. & Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. <i>Bioinformatics</i> 19:1572-1574.

1.5.2. Softwares

Um dos primeiros programas de computador desenvolvido especificamente para a análise de dados de genética de populações foi o *Biosys-1*. Programado em *Fortran* simples, ele executa as análises genéticas básicas (frequências gênicas, heterozigosidades, testes ao equilíbrio de Hardy-Weinberg, similaridade e distância gênica, estatística F de Wright etc.). Este programa foi extensivamente usado a partir da sua publicação, em 1981, quando, então, começou a dividir os seus méritos com programas mais recentes que calculam, também, as estatísticas básicas e descrevem a diversidade genética dentro e entre populações, como o *GENEPOP* e o *FSTAT*, ambos publicados em 1995 e o *Arlequim*, publicado em 2005.

Mais recentemente, têm surgido programas que utilizam a abordagem bayesiana. Esta abordagem analítica não é nova, contudo, devido ao grande número de iterações que demanda, só começou a se popularizar a partir do aumento do poder computacional. Nas abordagens bayesianas são obtidas probabilidades *a posteriori*, que são definidas em função de probabilidades *a priori*. As probabilidades *a priori* referem-se à distribuição dos valores do parâmetro antes da observação dos dados e, a probabilidade *a posteriori*, retrata a distribuição desses valores após considerar os dados. A distribuição *a posteriori* fornece, portanto, uma medida da incerteza relacionada às estimativas realizadas e, a sua média, pode ser usada como estimativa pontual. Nos métodos desenvolvidos mais recentemente, vários parâmetros populacionais são estimados a partir dos dados ao mesmo tempo. Assim, os cálculos da distribuição *a posteriori* são somas ou integrais complexas, que não podem ser resolvidas analiticamente e cuja aproximação é feita, geralmente, por simulação Monte Carlo com cadeia de Markov (MCMC, do inglês *Markov chain Monte Carlo*).

As MCMC modelam processos estocásticos com um número finito de estados nos quais o estado futuro depende apenas do estado presente do sistema e nunca do estado em momentos passados. Elas são usadas para obter a função de probabilidade posterior de diferentes genealogias (e parâmetros) dentre as possíveis, já que é impossível calcular todas elas devido à demanda computacional. A cadeia move-se pelo espaço de soluções possíveis, calculando a cada passo a distribuição posterior de uma mudança na genealogia e comparando-a com a atual. Se a distribuição posterior da nova genealogia for maior que a atual, a cadeia aceita a mudança. Dessa forma, são amostradas as genealogias que contribuem para estimativas mais prováveis dos parâmetros. Um dos pontos fortes de um método bayesiano feito por simulação MCMC é que, vários tipos de dados podem ser combinados em uma única análise e, também, vários parâmetros podem ser estimados simultaneamente. As análises bayesianas são poderosas e podem fornecer soluções quando outras abordagens não o conseguem.

Os programas mais utilizados para estimar, atualmente, a subdivisão populacional e construir filogenias utilizando análise bayesianas são o *Structure* e o *MrBayes*, respectivamente. O *Structure* usa inferência Bayesiana para, a partir dos genótipos multilocios dos indivíduos, estimar o número de populações que melhor explica os dados, ao mesmo tempo em que os indivíduos são atribuídos às populações. O critério para o agrupamento em populações é o de minimizar os desequilíbrios de Hardy-Weinberg e de ligação dentro delas. Assim, as populações são definidas pelos dados, sem qualquer inferência *a priori*. Posteriormente, a informação sobre a origem geográfica das amostras pode ser incorporada à análise para aumentar a resolução dos limites entre as populações. O *MrBayes* é um *software* que usa, também, inferências bayesianas, mas para escolha de modelos filogenéticos baseados em dados de sequências de DNA.

Embora a abordagem bayesiana seja muito atraente por incluir o máximo de informação e reconstruir eventos complexos, elas nem sempre são a melhor abordagem. Se os dados são poucos ou não informativos, a análise bayesiana não é adequada e, certamente, as inferências tradicionais trazem mais informação do que ela, sendo, portanto, muito mais adequadas nestes casos.

1.5.3. Em Frente para o Passado

Dada a velocidade e a facilidade com que podem ser recolhidos dados moleculares e as interfaces de fácil utilização dos *softwares*, quase qualquer um com um computador pode coletar e analisar dados e tentar publicar em Genética Marinha, independentemente da sua formação. O número de estudos explodiu durante as últimas décadas e, em função disso, alguns equívocos têm sido produzidos nessa massa de trabalhos científicos, alguns deles (embora não exclusivamente, ver os problemas do fetichismo das técnicas em **1.3.8. E Agora José?**) relacionados com a aplicação acrítica dos programas.

Muitos dos equívocos encontrados na produção científica em geral e em Genética Marinha em particular advêm da atitude: “é só formatar corretamente os dados para os programas que o resto está no manual”. Contudo, sem precisar da teoria da genética da população, uma vez que eles podem buscá-la ao longo do caminho ou que todas as informações que eles precisam “estão no manual”. Espera-se que toda a discussão travada até aqui, tenha convencido de que para trabalhar com Genética Marinha é fundamental a compreensão das premissas e limitações das diversas análises disponíveis. Fazer ciência é uma luta cotidiana para retificação de erros e isso só pode ser feito recolhendo dados em função de perguntas e, os resultados, avaliados de

maneira crítica (ver adiante na seção **1.6. Científico**). A figura 1.29 resume as duas atitudes com relação à análise de dados referidas aqui.

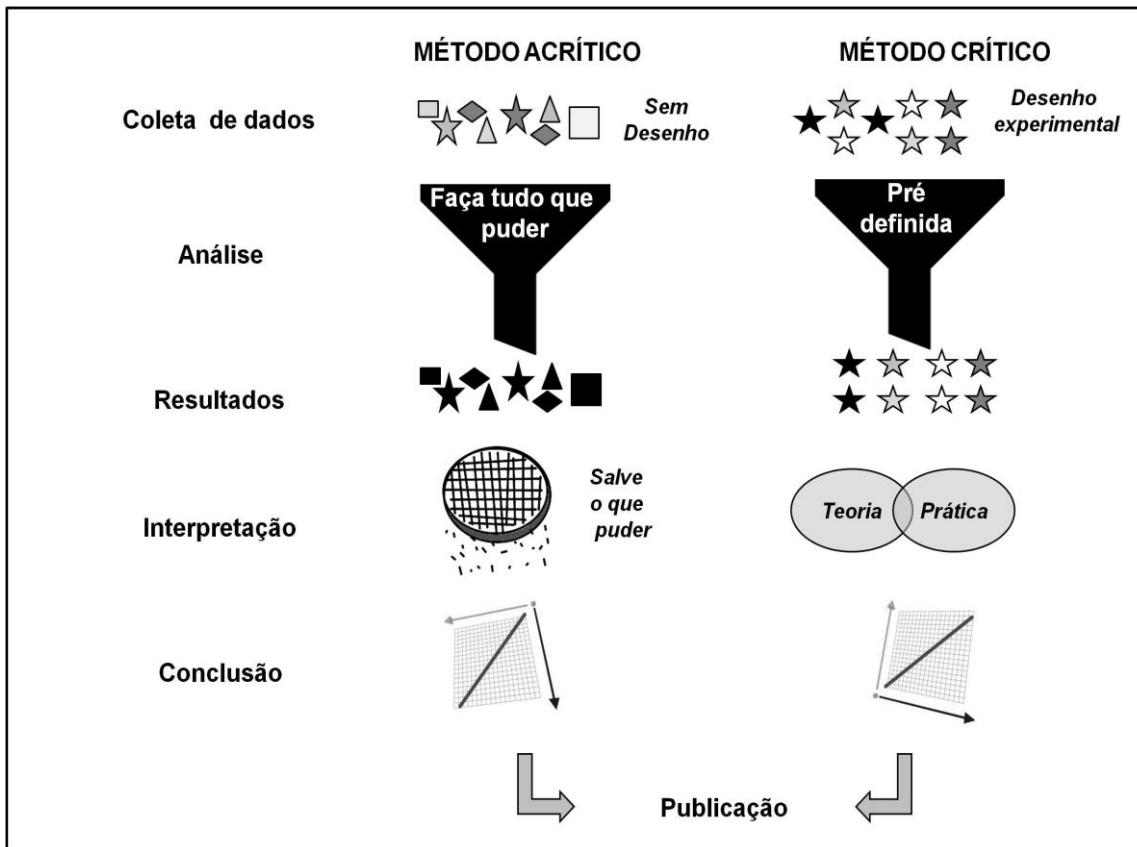


Figura 1.29. Representação irônica das atitudes relativas ao processo de produção científica como descritas no texto.

1.6. CIENTÍFICO

De posse das técnicas e métodos de amostragem da variação (gênica ou fenotípica) e das ferramentas conceituais de interrogação do mundo e interpretação das suas respostas, é chegada a hora de se fazer ciência. Mas o que é isso? Certamente uma forma de conhecimento, o tão incensado conhecimento “científico”. Mas o que significa este adjetivo? A ciência é filha da filosofia. Desligou-se da sua mãe e buscou seu próprio caminho por volta do século XIX. Nascia ali a ciência na sua versão moderna. Uma das mais fortes aliadas do capitalismo para aumentar a sua capacidade de produção e extração de mais valia (Loose 1993; Marcondes 2008).

O caminho que a ciência tomou, e que a separou da filosofia, foi aquele da construção de um objeto de estudo. Construção teórica, que seja bem entendido, uma vez que a atividade científica não se constitui, simplesmente, em tomar do mundo um objeto particular, mas de produzir, racionalmente, este objeto. É possível se imaginar, portanto, que a atividade científica é mais modesta do que aquela da filosofia, que se propõe a formular explicações gerais para o mundo, o universo, o ser etc. Ou seja, a filosofia se preocupa com o universal e a ciência com o particular. No entanto, é preferível distinguir mãe e filha pelas suas empreitadas: geral e universal para a filosofia, particular e específica para a ciência.

É possível, ainda, usar uma metáfora para caracterizar o trabalho da filosofia e da ciência. A filosofia teria um interesse horizontal (o geral, o universal) e a ciência vertical (o particular, o específico). Neste sentido, ambas as empreitadas são ilimitadas quanto ao movimento que executam na sua busca: não existem lados que limitem o conhecimento filosófico ou fundo que pare o conhecimento científico. O adjetivo “científico”, portanto, se define pelas características que definem a atividade científica.

Mas o quê na atividade científica a faz ser “científica”? A resposta para esta pergunta é: depende de quem está respondendo! Contudo, para não se ficar “em cima do

muro” é possível dar uma resposta que, embora não resolva todo problema nem seja unânime, sublinha uma característica típica da atividade científica. Esta resposta foi dada por um epistemólogo francês chamado Gaston Bachelard.

Segundo Bachelard, não se deve ter a ilusão que em ciência se trabalha diretamente com o real e, nem ao menos, que o conhecimento científico, de algum modo, se aproxima deste real. A questão do conhecimento científico é diferente e se refere à capacidade que o conhecimento (construído a partir dos seus métodos e ferramentas conceituais) tem de agir sobre a realidade, ou seja, o quanto realizável ele é. Por exemplo, toda vez que se constrói um experimento no laboratório, se está se afastando da realidade “real”. Isto porque nos experimentos científicos é retirado um monte de variáveis que poderiam atrapalhar a interpretação dos resultados. Ou seja, o mundo dentro do laboratório é, em certo sentido, um mundo inventado, construído para que perguntas possam ser feitas e, fundamentalmente, que seja possível se obter respostas para elas (Bachelard 1978; 1983).

Assim, o mundo do laboratório é outro mundo, diferente do mundo real. O laboratório é uma fenomenotécnica: fenômenos construídos para funcionar como se precisa que eles funcionem, geralmente, com todas as condições controladas, exceto uma, que é aquela que se quer testar. Todo mundo que trabalha ou já trabalhou em um laboratório sabe que, se fosse “a vera”, no mundo, seria muito mais difícil executar os experimentos. Portanto, o científico aqui apresentado deve ser entendido assim, como uma realidade científica construída a partir da utilização das técnicas e métodos de amostragem da variação (gênica ou fenotípica) e das ferramentas conceituais de interrogação do mundo e interpretação das suas respostas. Este é o científico da Genética Marinha a que se refere aqui.

1.6.1. Variação Gênica

O debate entre neutralistas e selecionistas (ver 1.2.1.3. *Controvérsias evolutivas*) guiou grande parte das pesquisas em Genética Marinha nas décadas de 1970 e 1980. Contudo, ainda hoje, o interesse em caracterizar os níveis de variação gênica das populações naturais de organismos marinhos não acabou, embora tenha se arrefecido bastante, especialmente diante da massa de dados que foi produzida nestes 50 anos de pesquisa. Organismos marinhos apresentam, geralmente, altos níveis de variação gênica (ver Quadro 1.21) (Solé-Cava & Thorpe 1989; 1991). No entanto, compreender quais são os mecanismos que mantém os níveis de variação observados é, ainda, contencioso.

Os selecionistas argumentam que os níveis de variação gênica são mantidos por seleção natural e os neutralistas defendem que a variação gênica é mantida por um equilíbrio entre taxa de mutação e tamanho efetivo de população (ver 1.2.1.3. *Controvérsias evolutivas*). A escala e a extensão que estes fatores determinam os níveis de variação gênica permanecem mal compreendidas e, provavelmente, variam, dependendo do organismo, sua estrutura populacional, da heterogeneidade de *habitat* e da estabilidade do ambiente. Apesar de não se encontrar mais como ponto central dos trabalhos de Genética Marinha, a quantidade de variação gênica nas populações naturais e seus mecanismos mantenedores são, ainda, um problema estudado e discutido.

Quadro 1.21. Níveis de variação gênica em diferentes grupos taxonômicos.

Variação gênica de grandes grupos e alguns grupos de organismos marinhos.

Grandes grupos	Grupos marinhos	Polimorfismo	Heterozigosidade
Plantas		45%	0,17 - 0,19 (Torggler et al. 1995)
	Plantas de mangue	18%	0,00 - 0,25 (Dodd & Rafii 2002)
Invertebrados		40 - 50%	0,12 - 0,15
	Bivalves	44%	0,17
	Gastrópodos	46%	0,12
	Esponjas	63%	0,27
	Cnidários	70%	0,25
Vertebrados		14 - 30%	0,03 - 0,08
	Peixes marinhos	29,1%	0,059

Exemplos de variação gênica (medidos como heterozigosidades médias) de alguns organismos marinhos para diferentes marcadores moleculares.

Grupos marinhos	Espécie	Marcador	Heterozigosidade
Plantas	<i>Avicennia</i> spp.	Aloenzimas	0,071 (Duke et al. 1998)
		AFLP	0,183 (Duke et al. 1998)
		Microsatélite	0,494 (Duke et al. 1998)
Gastrópodo	<i>Littorina saxatilis</i>	Aloenzimas	0,449
	<i>Littorina saxatilis</i>	Microsatélite	0,678
Bivalve	<i>Macoma balthica</i>	Aloenzimas	0,385 (Becquet et al. 2013)
	<i>Macoma balthica</i>	Microsatélite	0,538 (Becquet et al. 2013)
Crustáceo	<i>Balanus amphitrite</i>	Aloenzimas	0,103
	<i>Balanus amphitrite</i>	Microsatélite	0,845 (Robson et al. 2009)
Equinodermos	<i>Paracentrotus gaimardi</i>	Aloenzimas	0,232 (Duarte et al. 2016)
	<i>Paracentrotus lividus</i>	Microsatélite	0,910 (Calderón et al. 2009)
Peixes	<i>Sardinella</i> spp.	Aloenzimas	0,043
	<i>Theragra chalcogramma</i>	Microsatélite	0,850 (O'Reilly et al. 2004)

Na década de 1970, Jeffrey Levinton realizou um estudo para dois *loci* de aloenzimas (*Pgi* e *Lap*, ver Quadro 1.22) em oito espécies do gênero *Macoma* de moluscos bivalves, duas das quais da zona entremarés e seis espécies da zona sublitoral (Levinton 1975). O objetivo de Levinton era testar a hipótese selecionista de que a variação temporal do ambiente favorece os indivíduos heterozigotos e desta forma determina maiores níveis de variação gênica. Assim, as duas espécies de *Macoma* da zona entremarés deveriam apresentar níveis de variação gênica mais elevados do que as outras seis do sublitoral. Isto, contudo, não foi observado. Por outro lado, outro estudo,

também com aloenzimas, agora com o mexilhão *Perna canaliculus*, para as mesmas duas zonas litorâneas, encontrou heterozigosidades substancialmente maiores para os indivíduos da zona entremarés, de acordo com os pressupostos da hipótese selecionista de variação temporal no ambiente (Sin et al. 1990).

A lógica por trás da hipótese de que a variação temporal no ambiente é um mecanismo mantenedor de variação gênica nos organismos, se baseia na ideia de que, nestes casos, existe sobredominância, ou seja, admitindo-se que diferentes alelos apresentam diferentes respostas a determinados fatores ambientais, organismos heterozigotos seriam mais aptos a responder ao gradiente de variação ambiental do que os homozigotos.

Outra forma possível de explicar o fenômeno com base nos mesmos pressupostos é analisá-lo como sendo resultado da heterogeneidade espacial do ambiente. Neste caso, os níveis de variação gênica dependeriam do número de subdivisões percebidas pelo organismo no ambiente, o chamado “grão ambiental”. Assim, organismos bentônicos e sésseis são representantes típicos de uma estratégia de “grão grosso”, ou seja, de organismos que percebem inúmeras subdivisões no seu meio. Esta sensibilidade dos organismos às diferenças espaciais do ambiente pode ser relevante no momento do recrutamento (que pode acontecer em épocas e grupos discretos), quando as larvas podem efetuar uma seleção ativa de substrato e, também, na suscetibilidade destas larvas a tipos variados de seleção após o seu assentamento (Johnson & Black 1982; 1984a).

Naquele mesmo trabalho realizado com as espécies de *Macoma*, foi observado que a heterogeneidade ambiental pode selecionar alguns *loci* no sentido de um aumento do polimorfismo. Do mesmo modo, em outro estudo, agora amostrando 17 *loci* aloenzimáticos para o gênero de gastrópode *Littorina*, Eviatar Nevo e colaboradores

demonstraram, também, que altas heterozigosidades podem ser favorecidas em ambientes espacialmente mais heterogêneos (Noy et al. 1987). Contudo, em todos estes casos, o incremento de variação gênica das populações (polimorfismo, heterozigosidade) em ambientes heterogêneos não estava associado, obrigatoriamente, a um aumento da adaptabilidade dos heterozigotos. Isto é um problema para a hipótese de que a variação gênica está sendo mantida por seleção natural, uma vez que uma simples correlação entre variação gênica e heterogeneidade ambiental não define qual o mecanismo que está determinando esta correlação.

Quadro 1.22. As enzimas LAP e PGI (Hilbish & Koehn 1985a; 1985b; Garton & Berg 1989; Chan 1993).

As enzimas são catalizadores de origem biológica classificadas pela Comissão Internacional de Enzimas de acordo com as reações que catalisam. Elas recebem um nome sistemático e uma identificação de quatro dígitos, separados por pontos, precedida pela sigla EC (número de comissão). O primeiro número corresponde à classe a que a enzima pertence, o segundo determina a subclasse, o terceiro define o tipo de atividade enzimática e o quarto corresponde ao número da enzima dentro da sua subclasse.

Enzima LAP - Leucina Aminopeptidase (EC 3.4.11.1)

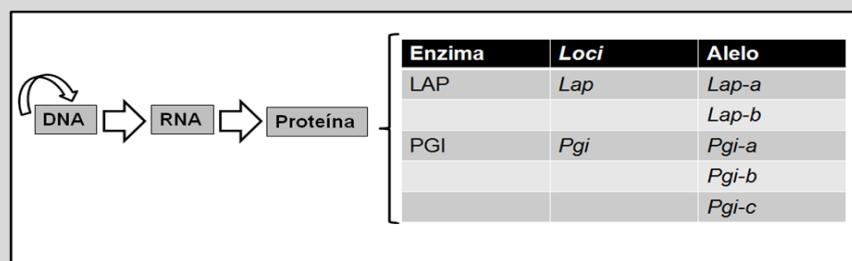
A leucina aminopeptidase é uma enzima constituída de quatro subunidades. Cada uma possui um átomo de Zinco, pertencente ao grupo das exopeptidases cuja atividade proteolítica é exercida sobre resíduos de leucina no aminoácido N-terminal dos peptídeos. Apesar de seu nome, a sua especificidade é fraca e os resíduos de outros aminoácidos N-terminal também podem ser clivados. Sua ação precisa da presença de íons metálicos bivalentes Mg ou Mn.

Enzima PGI - Fosfoglucose Isomerase (EC 5.3.1.9)

A fosfoglucose isomerase é uma enzima constituída de duas subunidades. Esta enzima catalisa a isomerização reversível da glicose 6-fosfato e da frutose 6-fosfato. Uma vez que a reação é reversível, a sua direção é determinada pelas concentrações do substrato e do produto.



Na amostragem da variação gênica (ver **1.3.5. Eletroforese de Aloenzimas**), as atividades catalíticas das enzimas são usadas para marcar um suporte (geralmente um gel de amido) com a atividade das suas diferentes isoformas. As isoformas são as enzimas de mesma função, mas com a sequência de aminoácidos diferente. A diferença na sequência de aminoácidos das isoformas tem origem em pequenas substituições de nucleotídeos no DNA responsável pela sua codificação. Estas substituições são grandes o suficiente para mudar a mobilidade da enzima em uma eletroforese (de forma que podem ser identificadas como isoformas da mesma enzima), mas pequenas o suficiente para não mudar a função da enzima, ou seja, sua atividade catalítica (de modo que as isoformas podem ser consideradas isoenzimas). Esta é a lógica que faz com quê as isoenzimas possam ser consideradas aloenzimas, ou seja, isoenzimas que apresentam um padrão mendeliano. Assim, aloenzimas podem ser consideradas diferentes alelos de um *locus* que é a função catalítica de uma enzima. Quando as diferentes enzimas são consideradas diferentes *loci* gênicos, elas são notadas pela sua sigla em itálico e com apenas a primeira letra em maiúscula (por exemplo, LAP-a enzima por *Lap-o locus*; PGI-a enzima por *Pgi-o locus*). As diferentes aloenzimas passam a identificar os diferentes alelos (ou alelomorfos) do *locus*.



Os vertebrados marinhos apresentam, geralmente, níveis de variação gênica mais baixos do que aqueles observados para invertebrados marinhos (ver o Quadro 1.21). Contudo, alguns *loci* de vertebrados se apresentam altamente variáveis. Este é o caso, por exemplo, das regiões codificadoras do complexo principal de histocompatibilidade ou MHC (do inglês *Major Histocompatibility Complex*) que desempenham um papel importante na resposta imunitária e estão entre as regiões codificadoras mais polimórficas conhecidas. O *locus* de MHC Classe II DQB foi estudado para uma população migratória de baleias azuis (*Balaenoptera musculus*) que ocorre no Golfo da Califórnia. Um fragmento de 172bp do *locus* foi sequenciado para 80 indivíduos desta população e os resultados encontrados sugeriram que a seleção natural agiu sobre este *locus* ao longo de períodos prolongados de tempo (Moreno-Santillán et al. 2016).

Em um estudo mais extenso sobre variação gênica de vertebrados marinhos, Alicia Dalongeville (a época uma estudante de doutorado da Universidade de Montpellier, na França), seus orientadores e colaboradores levantaram dados de marcadores moleculares do tipo microsatélite para 13 espécies de peixes do Mar Mediterrâneo. O objetivo era tentar correlacionar os padrões de variação gênica com caracteres ecológicos tais quais distribuição (vertical e horizontal), migração, estratégia reprodutiva e comprimento do corpo das espécies. Os resultados evidenciaram que alguns dos fatores ecológicos investigados eram capazes de explicar (ou seja, eram correlacionados de forma estatisticamente significativa) uma proporção importante da diversidade genética encontrada nos *loci* de microsatélite, sugerindo que as estratégias de vida das espécies podem influenciar a variação gênica. Por exemplo, mais da metade da variação gênica encontrada para as 13 espécies estudadas pôde ser explicada pelos fatores ecológicos distribuição vertical (variação na profundidade), migração e comprimento do corpo. A distribuição vertical mostrou uma associação significativa

com a heterozigosidade. Do mesmo modo, as espécies migratórias apresentaram maior variação gênica do que as não migratórias. A distribuição horizontal (distribuição longitudinal), no entanto, não apresentou correlação com a variação gênica (Dalongeville et al. 2016).

Em todos os casos descritos até aqui, o objetivo explícito dos autores foi tentar correlacionar os níveis de variação gênica encontrados com o ambiente, ou seja, tentar atribuir à seleção natural o papel de mecanismos responsável pela origem e manutenção dos níveis de variação gênica observados. Contudo, muitos estudos sobre variação gênica em populações naturais de organismos marinhos não consideram a seleção natural. Muito pelo contrário, a variação observada é considerada neutra, ou seja, sem valor seletivo algum. Nestes casos, a explicação segue os pressupostos neutralistas, sendo a variação gênica mantida por um equilíbrio entre o tamanho efetivo da população e a taxa de mutação (ver Quadro 1.23). Por exemplo, os níveis de variação gênica nos moluscos bivalves *Anomalocardia brasiliiana* (conhecido popularmente como berbigão, marisco pedra, vongôle) e *Perna perna* (conhecido popularmente como mexilhão marron, marisco), ambos com ocorrência no litoral brasileiro, foram amostrados com o método de eletroforese de aloenzimas. O mexilhão *Perna perna* se mostrou com o dobro da variação gênica daquela observada para o berbigão *Anomalocardia brasiliiana*. Ambos os organismos habitam a zona entremarés, contudo, o mexilhão vive em costões rochosos e o berbigão em praias arenoso-lodosas. A explicação para os níveis de variação gênica encontrados poderia, facilmente, ser buscada nas diferenças ambientais a que estão sujeitos estes dois organismos. Contudo, neste caso, de maneira simples, a explicação para as diferenças de variação gênica encontradas se baseou apenas na inferência dos tamanhos de população destas duas espécies (ver Quadro 1.24).

David Skibinski e seus colaboradores utilizando-se de um grande banco de dados contendo informações sobre variação gênica de aloenzimas para mais de 1500 espécies (de todos os ambientes e *taxa* e não só marinhas), testou as expectativas teóricas das hipóteses neutralista e selecionista em relação às heterozigosidades encontradas. A conclusão foi que os resultados forneciam suporte para a hipótese de que uma grande percentagem da variação gênica observada podia ser explicada pela variação na taxa de mutação neutra e, uma percentagem menor, pela ação de seleção natural (Skibinski et al. 1983). Portanto, por um lado... mas pelo outro...

Quadro 1.23. Explicação neutralista para variação gênica e definição de tamanho efetivo de população.

HETEROZIGOSIDADE ESPERADA PARA ALELOS NEUTROS

Se a variação alélica for neutra, o padrão de variação gênica das populações é determinado por um balanço entre a deriva genética e taxa de mutação. Desta forma, a teoria neutralista prevê que a heterozigosidade pode ser estimada a partir da equação:

$$H = \frac{4Ne\bar{v}}{1 + 4Ne\bar{v}}$$

Onde:

H = Heterozigosidade

Ne = Tamanho efetivo de população

\bar{v} = Taxa de mutação

Esta equação evidencia que para populações com grandes tamanhos efetivos são esperadas altas heterozigosidades. De modo inverso, para populações de pequeno tamanho efetivo são esperadas baixas heterozigosidades (rever a Figura 1.4 seção 4.2.1.3. *Controvérsias evolutivas*).

TAMANHO EFETIVO DE POPULAÇÃO

O tamanho nominal de uma população (N) é diferente do tamanho efetivo de população (Ne). O tamanho nominal diz respeito ao censo do número de indivíduos de uma população. O tamanho efetivo, por outro lado, é definido como o tamanho de uma população ideal segundo os pressupostos do modelo Wright-Fisher. Neste modelo, os indivíduos são diplóides, as gerações são discretas (não sobrepostas), a migração garante a homogeneidade das frequências e cada alelo é herdado da geração anterior de modo independente e aleatório.

O tamanho efetivo de população, portanto, leva em consideração apenas os indivíduos que participam efetivamente da reprodução. Indivíduos muito jovens ou muito velhos para reprodução não contam. Do mesmo modo, o tamanho efetivo de população é sensível a mudanças temporais do número de indivíduos que participam da reprodução, como, também, a desvios das proporções sexuais.

Em função das proporções sexuais, o tamanho efetivo de população pode ser calculado usando a equação:

$$Ne = \frac{4NmNf}{(Nm + Nf)}$$

Onde:

Ne = Tamanho efetivo de população

Nm = Número de machos

Nf = Número de fêmeas

Assim, se forem consideradas três populações, ambas com tamanho nominal igual a 100 indivíduos, mas com diferentes proporções sexuais, os tamanhos efetivos seriam como mostrados abaixo:

Populações	N	Nm	Nf	Ne
1	100	50	50	100
2	100	10	90	36
3	100	1	99	4

Quadro 1.24. Estimativa dos tamanhos de população para *Anomalocardia brasiliiana* e *Perna perna* (Solé-Cava et al. 1992; Silva & Solé-Cava 1994).

Como visto no quadro 1.23 a hipótese neutralista é capaz de prever, matematicamente, a heterozigosidade de uma população com base no tamanho efetivo de população (N_e) e taxa de mutação (v):

$$H = \frac{4Nev}{1 + 4Nev} \quad (\text{equação 1})$$

Com um pouco de álgebra, é possível transformar a equação 1 em:

$$Nev = \frac{H}{(4 - 4H)} \quad (\text{equação 2})$$

Dadas as heterozigosidades observadas para as espécies *Anomalocardia brasiliiana* e *Perna perna* é possível, usando a equação 2, estimar, de maneira simples, o $N_e v$ para as populações estudadas destas duas espécies:

Espécie	Heterozigosidade	Heterozigosidade Média	Nev
<i>A. brasiliiana</i>	0,16-0,26	0,21	0,07
<i>P. perna</i>	0,32-0,43	0,38	0,15

Neste ponto, é interessante tentar resumir as hipóteses neutralista e selecionistas (e suas variantes) para explicar a variação gênica. Este resumo se encontra na figura 1.30 e, como pode ser percebido, algumas previsões selecionistas podem ser explicadas, também, por previsões neutralistas. Este é o caso, por exemplo, das relações entre tamanho do corpo e variação gênica e homeostase e variação gênica. É fácil perceber que organismos maiores são, também, aqueles que apresentam menores tamanhos de população, portanto, menor variação gênica. Do mesmo modo, baixa homeostase é típica de organismos mais antigos na evolução (como poríferos e cnidários) e que, também, apresentam maiores tamanhos de população que aqueles organismos de alta homeostase, como os vertebrados. A persistência da espécie no ambiente e a sua idade filogenética, também podem ser consideradas tanto de um ponto de vista neutralista quanto selecionista, uma vez que podem ser interpretadas como resultado de estabilidade demográfica ou adaptação. Ou seja, como foi dito no começo desta seção, a escala e a extensão que a variação gênica é determinada por seleção natural (hipóteses

selecionistas) ou é resultado da ação de deriva genética sobre alelos neutros (hipótese neutralista) deve estar na dependência do organismo, dos *loci*, da estrutura populacional da espécie, da estabilidade do ambiente etc.

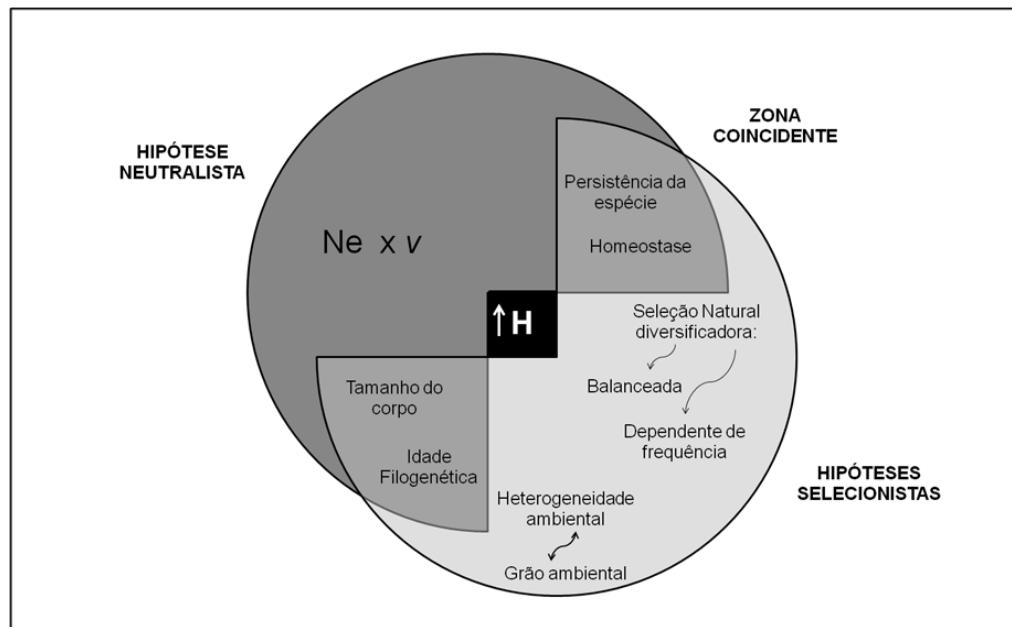


Figura 1.30. Representação esquemática das explicações para os altos níveis de variação gênica relacionadas às hipóteses neutralista e selecionista (Ne = tamanho efetivo de população; v = taxa de mutação e $H\uparrow$ = aumento da heterozigosidade). Maiores explicações sobre as hipóteses específicas podem ser encontradas no texto desta seção, como também em 1.2.1.3. *Controvérsias evolutivas*, especialmente a figura 1.4 no que diz respeito à hipótese neutralista.

A Convenção sobre Diversidade Biológica das Nações Unidas que aconteceu no Rio de Janeiro no ano de 1992 (*Convention on Biological Diversity, Rio de Janeiro, Brazil*), esteve preocupada com a redução da diversidade biológica devido a atividades humanas e definiu a variação gênica como um dos três níveis de diversidade biológica (juntamente com variação de espécies e de ecossistemas) a ser preservado, sendo ela essencial para a persistência das espécies em ambientes em mudança (United Nations 1992). Ambientes como os manguezais são, por exemplo, alvo de grande preocupação, uma vez que são afetados tanto por causas naturais (como oscilações do nível do mar) quanto pelas atividades humanas. Muitos *taxa* de mangue possuem ampla distribuição

geográfica, que tem sido atribuída a uma eficiente distribuição de propágulos que, provavelmente, proporciona uma coesão genética entre as populações. No entanto, como observado em outros organismos marinhos, estudos de genética com espécies de mangue sugerem que o fluxo de genes é menos eficaz do que o esperado. Desta forma, as populações marginais, geralmente exibem níveis mais baixos de variação gênica em comparação com populações centrais, como resultado de deriva genética ou da pressão de seleção natural. Assim, além da simples informação sobre os níveis variação gênica, a Genética Marinha tem que tentar compreender os padrões de estruturação gênica das espécies marinhas e suas populações, o que será abordado a seguir.

1.6.2. Estruturação

As espécies, geralmente, não se distribuem homogeneamente por amplas extensões geográficas, mas são divididas em um número variável de populações. Estas populações podem ser mais ou menos semelhantes, dependendo de vários fatores tais quais os tamanhos destas populações (suscetibilidade aos efeitos de deriva genética), o quanto sua constituição genética é influenciada pelo ambiente do local que habitam (ação da seleção natural), o quanto estão conectadas entre si (efetividade da migração) etc (Slatkin 1985; Hedgecock 1986; Parker et al. 1998). O nível de diferenciação genética encontrado entre estas populações é o que se define por estruturação. Ou seja, absoluta homogeneidade genética significa ausência de estruturação. Por outro lado, à medida que as espécies vão se diferenciando pela sua suscetibilidade às circunstâncias, elas vão se tornando mais e mais estruturadas, ou seja, subdivididas em populações ou com populações diferenciadas. Os estudos de estruturação podem ser realizados através de marcação e recaptura dos indivíduos, descrição (ou etiquetagem) química e, o que interessa aqui, por intermédio de marcadores genéticos.

O sucesso na dispersão deixa um rastro genético que revela o padrão de conectividade genética. O grau de diferenciação entre as populações são estimados através de índices de identidade gênica e índices de estruturação (F_{ST}) (ver **1.4.2. Variação Interpopulacional**). Os diferentes marcadores genéticos são capazes de evidenciar dispersão em diferentes escalas de tempo. Os marcadores de frequência, como os microssatélites e as aloenzimas, são mais efetivos em evidenciar diferenciação em tempos ecológicos ou contemporâneos. Os marcadores de sequência como mtDNA e SNP, são indicados para tempos maiores ou históricos (ver **1.3.5. Eletroforese de Aloenzimas e 1.3.6. Técnicas de DNA**).

O ambiente marinho aparenta ser homogêneo e, devido a este fato, espera-se que as populações de organismos marinhos não sejam estruturadas, pois, mesmo os organismos sésseis, podem passar parte da sua vida (fase de vida larvar) no plâncton, o que os permite dispersar por grandes distâncias. Entretanto, as populações marinhas não estão completamente conectadas, mas apresentam diferentes padrões de estruturação. Os principais mecanismos utilizados para explicar a diferenciação das populações nos ambientes marinhos são os padrões de correntes, gradientes ambientais (salinidade, temperatura etc), seleção pós-assentamento e ecologia do adulto.

Os organismos com maior capacidade de dispersão, como os que possuem larva planctotrófica, apresentam, geralmente, populações com altos índices de identidade gênica e baixos índices de estruturação (F_{ST}). Em contrapartida, os organismos com menor capacidade de dispersão, como aqueles que são sésseis ou apresentam larvas lecitotróficas, apresentam, geralmente, populações com baixos índices de identidade gênica e altos índices de estruturação. Ou seja, a expectativa é regida pelo modelo de isolamento por distância, porém, não raro, ela pode ser frustrada.

Diferente das expectativas, a dispersão efetiva de larvas e adultos é bem menor que a dispersão potencial. Portanto, o fluxo gênico que se realiza é, também, menor que a capacidade de dispersão dos organismos. Desta forma, a diferenciação geográfica dos organismos marinhos é bem maior do que à primeira vista poderia sugerir o seu modo de dispersão. Entre os fatores que determinam a diminuição do fluxo gênico entre populações de organismos marinhos estão, num primeiro momento, as barreiras como a salinidade, temperatura, disponibilidade de espaço e predação e, num segundo momento, a seleção pós-assentamento. Contudo, para começar, serão exemplificados alguns casos nos quais as expectativas do modelo de isolamento por distância não são frustradas.

Crassostrea gigas e *Saccostrea commercialis* são duas espécies comerciais de ostras com larvas planctotróficas que apresentam baixa estruturação (Identidade gênica = 0,99), mesmo em faixas geográficas amplas (> 500 Km) (Buroker et al. 1979a; 1979b). Outros invertebrados marinhos sedentários na vida adulta, mas com larvas planctotróficas, como o ouriço-do-mar *Echinometra mathaei* ($F_{ST} = 0,013$) e a estrela-do-mar *Acanthaster planci* ($F_{ST} = 0,070$), também são exemplos de espécies com alta capacidade de dispersão que apresentam, como esperado, homogeneidade genética entre as suas populações. No quadro 1.25 são apresentados alguns outros exemplos de organismos marinhos para os quais a diferenciação geográfica é função da capacidade dispersora.

Quadro 1.25. Níveis de diferenciação geográfica em organismos marinhos.

Exemplos de diferenciação geográfica em alguns organismos marinhos em função da capacidade dispersora da espécie. A diferenciação geográfica foi definida como pequena, moderada ou alta em função dos valores de F_{ST} e identidades gênicas descritos nos artigos científicos.

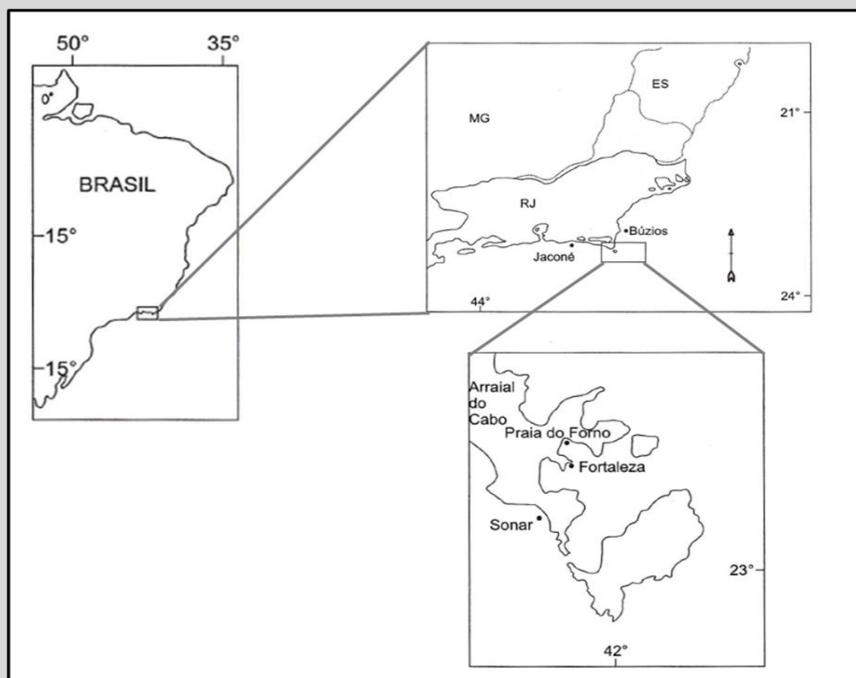
Organismo	Capacidade de dispersão	Diferenciação geográfica	Distância entre as populações (Km)
<i>Homarus</i> (lagosta)	Alta	Moderada	1 000
<i>Littorina</i> (gastrópodo)	Baixa	Alta	1
<i>Tigriopus</i> (copépodo)	Baixa	Alta	2
<i>Arbacia</i> (ouriço do mar)	Alta	Alta	1 000 (Marcus 1977)
<i>Ostrea</i> (ostra)	Alta	Pequena	2 000 - 3 000
<i>Crassostrea</i> (ostra)	Alta	Pequena	2 000 - 3 000
<i>Embiotoca jacksoni</i> (peixe)	Nula	Alta	300
<i>Clinocottus analis</i> (peixe)	Baixa	Moderada	300
<i>Alloclinus holderi</i> (peixe)	Limitada	Baixa	300

Os mexilhões, que são sésseis quando adultos, mas apresentam uma fase larvar planctotrófica, também são usados como exemplo do poder da dispersão larvar para homogeneizar populações distantes muitos quilômetros ao longo da costa. Contudo, um estudo realizado com populações do mexilhão *Perna perna* na região de ressurgência de Arraial do Cabo (Rio de Janeiro, Brasil) mostrou que o fluxo gênico entre populações do mexilhão distantes apenas 50 km era pequeno, quando comparado com populações da mesma espécie de outras regiões e com outras espécies de mexilhão. A explicação para estes dados parece estar relacionada ao fenômeno da ressurgência, já que a maior diferenciação foi observada, sempre, entre populações de regiões sob forte influência deste fenômeno e aquelas onde a ressurgência era ausente ou pouco influente. Desta forma, é possível dizer que um fenômeno oceanográfico como a ressurgência, que

provoca alterações abióticas (temperatura, salinidade etc.) e bióticas (produtividade, predação etc.) na massa d'água, pode estar funcionando como uma barreira geográfica nos mares e influenciando decisivamente a diferenciação entre populações de uma mesma espécie (ver Quadro 1.26).

Quadro 1.26. Estruturação do mexilhão *Perna perna* na região de ressurgência em Arraial do Cabo (Rio de Janeiro, Brasil).

Foram estudadas 5 populações (Forno, Fortaleza, Sonar, Búzios e Jaconé) da espécie do mexilhão *Perna perna* na e em torno da região de ressurgência de Arraial do Cabo. A localização destas cinco populações está mostrada no mapa.



A tabela abaixo mostra a influência da ressurgência para cada uma das cinco populações do mexilhão *Perna perna*, bem como outras variáveis que poderiam ser úteis para o estudo como, por exemplo, se o regime de ondas no local era agitado ou calmo (o chamado “modo”). Além da distância de cada população em relação ao “Sonar”, que foi escolhido como ponto zero, uma vez que é ali onde a maior intensidade de ressurgência é observada.

Local	Latitude	Longitude	Modo	Distância do Sonar (Km)	Ressurgência
Sonar	23°00'	42°00'	Agitado	0	Forte
Fortaleza	22°58'	42°00'	Agitado	1,5	Fraca
Forno	22°58'	42°00'	Calmo	2	Fraca
Búzios	22°46'	41°50'	Calmo	50	Ausente
Jaconé	22°54'	42°30'	Agitado	50	Forte

O índice de estruturação foi calculado pelo F_{ST} e, os resultados, comparados com índices de estruturação para outras populações de mexilhão no mundo. Fica claro, observando os resultados, que na região de ressurgência de Arraial do Cabo, os índices de estruturação são obtidos para 50 km de distância são equivalentes a distâncias 150 vezes maiores.

Espécie	Localização	Distância entre as populações (Km)	F_{ST}
<i>Mytilus edulis</i>	Reino Unido e Irlanda	1 200	0,0004
<i>Perna canaliculus</i>	Nova Zelândia	1 800	0,0868 (Gardner et al. 1996)
<i>Perna perna</i>	Africa do Sul	3 700	0,0333
<i>Perna perna</i>	Brasil	50	0,0752

Ainda com mexilhões, vários estudos com a espécie *Mytilus edulis* (o chamado mexilhão azul europeu) têm demonstrado variações nas frequências gênicas dos alelos da enzima *Lap* (ver Quadro 1.22) em distâncias também muito pequenas (50 km). A alteração na frequência de um dos alelos deste *locus* (*Lap*⁹⁴) acompanha um gradiente de salinidade que vai do oceano, onde a frequência do alelo alcança 55%, até o estuário onde a frequência do mesmo alelo cai para 15% (Koehn et al. 1976; 1980; Koehn 1978; 1985). Este tipo de correlação entre um fator ambiental e uma característica genética, associada a uma distribuição contínua no espaço, define um cline.

A explicação para este fenômeno pode ser resumida como sendo o resultado da ação intensa de seleção natural contra o alelo *Lap*⁹⁴ no ambiente estuarino. O mecanismo bioquímico de seleção parece estar associado à ação da *Lap* na mobilização de aminoácidos que regulam o volume celular (Hilbish et al. 1982). Indivíduos portadores do alelo *Lap*⁹⁴ apresentam taxas de excreção duas vezes maiores que aqueles dos genótipos constituídos pelos outros alelos. Em ambientes com flutuação nas condições de salinidade, como é o caso dos estuários, a presença de um alelo muito eficiente na excreção de aminoácidos acarreta um custo ecológico muito alto, levando a uma alta mortalidade. Desta forma, em ambientes estuarinos indivíduos com genótipos constituídos pelo alelo *Lap*⁹⁴ apresentam uma desvantagem. Nas condições oceânicas, onde existe maior estabilidade das condições de salinidade, o alelo *Lap*⁹⁴ não tem esta desvantagem e apresenta a maior freqüência (55%) em relação aos outros três alelos.

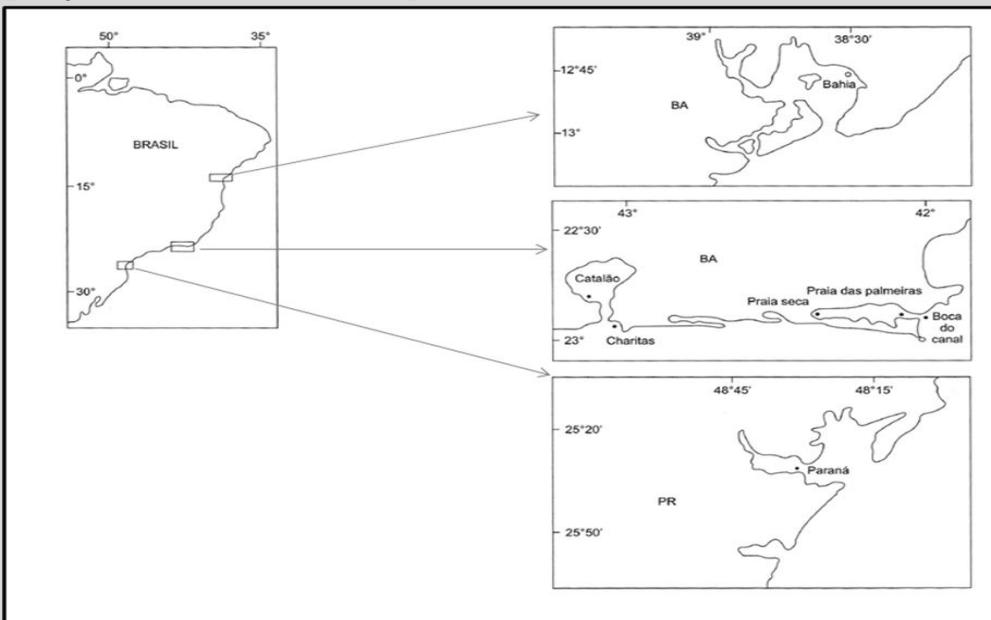
Este mecanismo bioquímico de regulação isosmótica no qual os genótipos do *locus Lap* estão envolvidos, é comumente encontrado em bivalves eurihalinos e já foram descritos fisiologicamente, embora não genéticamente, para outras espécies de moluscos bivalves como, por exemplo, *Mercenaria mercenaria*, *Callista chione*, *Mya*

arenaria, *Corbicula japonica*, *Modiolus demissus*, *Crassostrea virginica* e *Macoma inconspicua*, entre outros.

No Brasil, o polimorfismo da *Lap* foi estudado para o molusco bivalve eurialino *Anomalocardia brasiliiana*. Os resultados, também neste caso, indicaram a presença de diferenças significativas entre a freqüência dos alelos da *Lap* em função da salinidade. Os resultados não foram tão expressivos como aqueles descritos para a espécie *M. edulis* (ver Quadro 1.27), contudo, fortalecem a idéia de que alterações nas frequências gênicas da *Lap* são influenciadas fortemente pelas variações de salinidade. Ou seja, a seleção natural pode ter influência importante na diferenciação geográfica de populações, mesmo em distâncias relativamente pequenas e sem barreira geográfica aparente (Silva 2013).

Quadro 1.27. Diferenciação geográfica das frequências gênicas dos alelos do *locus Lap* em populações do molusco bivalve eurihalino *Anomalocardia brasiliiana* na costa brasileira.

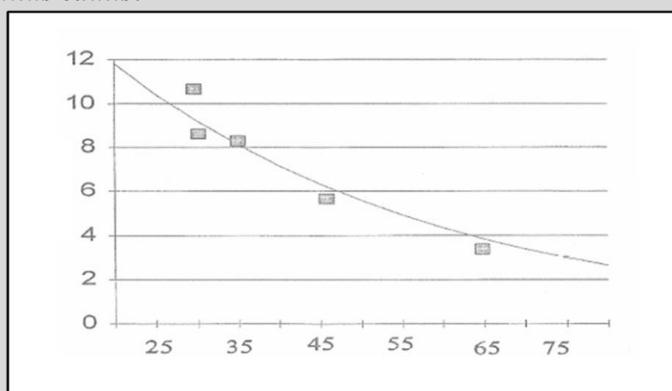
Sete populações da espécie *Anomalocardia brasiliiana* foram estudadas em quatro estuários ao longo 2000 km da costa brasileira. As localizações das populações estão mostradas no mapa abaixo.



As salinidades em cada um dos pontos, bem como a sua distância em relação ao ponto mais ao norte (Bahia) são dadas na tabela abaixo.

Local	Latitude	Longitude	Salinidade (%)	Distância da Bahia (Km)
Bahia	12°55'	38°29'	32,5	0
Boca do Canal	22°53'	42°00'	35,0	1 410
Palmeiras	22°52'	42°03'	45,0	1 420
Praia Seca	22°54'	42°20'	65,0	1 430
Charitas	22°58'	43°10'	32,0 – 36,9	1 550
Catalão	22°50'	43°15'	24,0 – 32,0	1 570
Paraná	25°30'	48°33'	28,0 – 31,0	1 950

A figura abaixo mostra a correlação entre a frequência de um dos alelos (*Lap^C*) com a salinidade. A correlação é significativa ($p < 0,003$) em condições hipoalinas, embora não para condições de altas salinidade. Estes resultados estão de acordo com aqueles descritos para espécie *Mytilus edulis*.



Todos os exemplos mencionados tentaram deixar claro como o processo de estruturação pode estar se dando no mar, a despeito da impressão de alta homogeneidade deste ambiente e da alta capacidade dispersiva da sua biota (Berger, 1973; Grant et al. 1992; Beaumont & Nichols 1996; Allendorf & Seeb 2000). Contudo, a história não acaba aí. Tudo que foi descrito até aqui disse respeito a casos que foram estudados com base em dados de aloenzimas. Estudos sobre estruturação baseados em marcadores de DNA ou, mais que isso, comparando resultados de aloenzimas e DNA, têm trazido resultados interessantes, senão inesperados.

Um estudo com corais do gênero *Pecillopora* em quatro ilhas no Pacífico sul (Moorea, Tahiti, Tetiaroa e Maiao) as quais, além de estarem separadas por pequenas distâncias (3 a 130 km), são atravessadas por uma mesma corrente, evidenciou que os corais de cada ilha eram completamente diferenciados entre si. A estruturação, neste caso, foi inferida a partir de dados de mtDNA. A explicação para este nível de diferenciação genética entre populações tão próximas geograficamente e contando com um veículo de migração tão eficiente como uma corrente, está na forma de recrutamento. Como uma espécie de coral, o recrutamento é altamente local devido à captura das larvas nas piscinas de maré que são comuns neste tipo de ambiente (Edmunds et al. 2016).

Outro exemplo. A espécie de gastrópodo marinho bento-pelágico *Concholepas concholepas* é endêmica do sudeste do Pacífico, com uma ampla área de distribuição geográfica (de ambientes tropicais até ambiente subantárticos). A espécie possui uma fase larvar que pode durar até três meses na coluna de água, sugerindo que a espécie possui uma grande capacidade dispersiva e, portanto, deve apresentar uma baixa estruturação na sua área de distribuição seguindo, provavelmente, o modelo de isolamento por distância. Para testar esta expectativa, foram usados marcadores de

microsatélites associados a uma técnica de amostragem hierárquica. Assim, foram estudadas três regiões geograficamente distantes (Antofagasta Bay, Valdivia e Patagônia) e, para cada uma delas foram amostrados, também, pontos próximos geograficamente, contudo, que apresentavam diferenças ambientais. Para as três regiões foi observado, conforme o esperado, um padrão de isolamento por distância. Contudo, dentro das regiões foram observados diferentes padrões de estruturação. Dentro das regiões de Antofagasta Bay e Valdivia havia evidência de substancial fluxo gênico entre as populações, enquanto que para região da Patagônia, as populações se apresentaram altamente estruturadas. Os resultados indicaram, portanto, a existência de diferentes padrões de estruturação genética na dependência de diferentes fatores como, por exemplo, a topografia costeira, que era uma das variáveis estudadas (Cardenas et al. 2016).

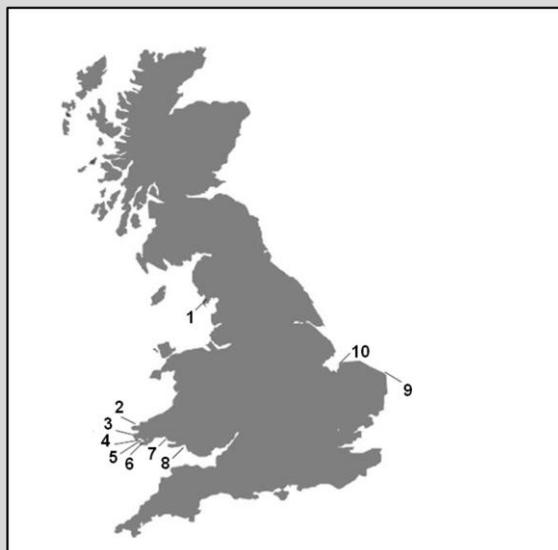
A partir da década 1990, os padrões de diferenciação populacional inferidos a partir de diferentes marcadores moleculares (aloenzimas, mtDNA, nDNA) começaram a ser comparados. Os resultados causaram surpresa, uma vez que diferentes padrões eram mostrados por diferentes marcadores, por exemplo, a ostra *Crassostrea virginica* que por aloenzimas mostrava populações geneticamente homogêneas por longas distâncias na costa atlântica dos EUA, se mostrava estruturada quando eram usados marcadores tanto de mtDNA quanto de nDNA (McDonald et al. 1996). A explicação para essa diferença, obviamente, teria que estar associada à dinâmica evolutiva a qual estes marcadores moleculares estariam sujeitos. Dito de outra forma, as forças evolutivas que estariam moldando o padrão de variação gênica nestes marcadores. Migração era o que estava se tentando inferir e a taxa de mutação é muito baixa, portanto, as forças evolutivas responsáveis pela diferença teriam que ser deriva genética e seleção natural. Uma delas estaria atuando em um dos conjuntos de marcadores e, a outra, no outro.

Como aloenzimas são proteínas e, portanto, têm uma função definida no organismo, a explicação para discordância entre padrões de diferenciação genética mostrados por aloenzimas e DNA teria que ser que as aloenzimas estariam sob seleção natural e o DNA seria neutro.

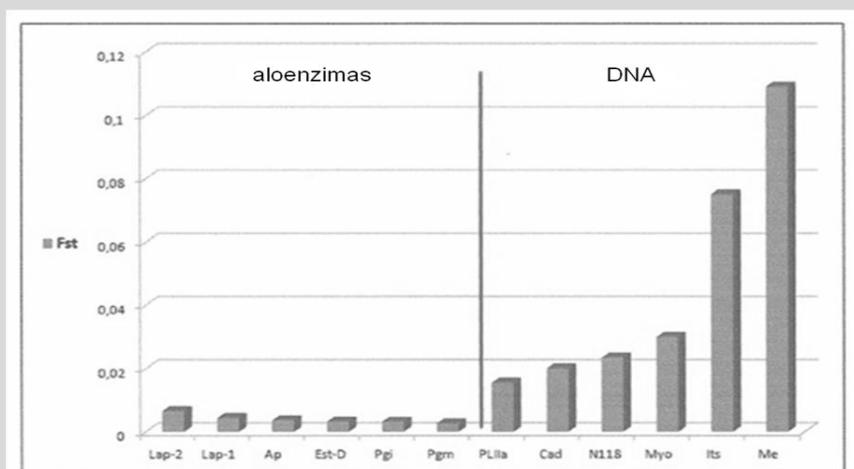
Além das ostras, estudos de estruturação genética com o copépodo *Tigriopus californicus* (Burton & Lee 1994) e com o peixe *Gadus morhua* (Pogson et al. 1995), apenas para citar os primeiros e mais impactantes trabalhos neste sentido, também mostraram discordância de padrões entre aloenzimas e DNA. A explicação escolhida por todos eles foi sempre a mesma. As aloenzimas estariam sob ação da seleção natural e, os marcadores de DNA, seriam neutros. Para voltar aos mexilhões, a espécie *Mytilus edulis* é estudada há décadas quanto a estruturação das suas populações ao redor da Grã-Bretanha e serviram, também, a um estudo comparativo dos marcadores moleculares de aloenzimas e DNA. Os resultados indicaram aquilo que já se estabeleceu: discrepâncias entre as estimativas de estruturação dadas por aloenzimas e DNA (Silva & Skibinski 2009). A explicação também foi a mesma: seleção natural atuando nas aloenzimas. Um resumo deste estudo se encontra no quadro 1.28.

Quadro 1.28. Estudo comparativo (aloenzimas e DNA) sobre a estruturação genética das populações de *Mytilus edulis* na costa da Grã-Bretanha.

Dez populações do mexilhão azul *Mytilus edulis* foram amostradas ao redor da Grã-Bretanha (1- Ravenglass, 2- Marloes Sands, 3- Milford Haven 2, 4- Dale Point, 5- Milford Haven 1, 6- Monkton Point, 7- Burry Port, 8- Swansea Bay, 9- Hunstanton, 10- The Wash) como mostrado no mapa abaixo.



A estruturação foi inferida a partir de análise de seis *loci* de nDNA. Os resultados foram comparados com dados históricos de estruturação gênica do mexilhão *M. edulis* obtidos de mais de 30 anos de estudos com aloenzimas. A figura abaixo mostra de maneira clara a diferença entre os dois marcadores. Os F_{ST} 's são muito maiores para os marcadores de DNA. Esta diferença visual é estatisticamente significativa (Teste U de Mann-Whitney). A conclusão é que as aloenzimas estão sob seleção natural.



O processo de estruturação pode manter, em diferentes regiões geográficas, populações nas quais a frequência gênica de determinados alelos é significativamente diferente e estável. Nesta situação, estas populações podem ser reconhecidas como raças geográficas. Se este processo se mantiver no tempo, pode resultar em uma diferenciação

mais extrema. As populações geográficas (raças) podem, por exemplo, apresentar diferentes alelos fixados. Neste caso, indivíduos anônimos quanto a sua origem geográfica podem ser identificados genéticamente como pertencendo à região de onde aquele alelo fixado é característico. Estes são os fatos e a lógica aplicada pela Genética Marinha quando aplicada a estudos de sistemática e taxonomia.

1.6.3. Sistemática e Taxonomia

O máximo da estruturação genética é a divisão das populações em espécies diferentes. Contudo, a definição do que vem a ser uma espécie não é consensual. De fato, existem, na atualidade, mais de 20 conceitos de espécie. O conceito mais popular e influente é, porém, o conceito biológico de espécie (CBE). Mayr definiu-o em seu livro *Populações, espécies e evolução* como: “Espécies são grupos de populações naturais intercruzantes, que são isoladas reprodutivamente de outros grupos intercruzantes de populações” (Mayr 1963). A idéia central no CBE é a de isolamento reprodutivo que deve ser primariamente intrínseco (genético) ao invés de extrínseco (barreira geográfica). As espécies são vistas como sistemas fechados, ou seja, sem fluxo gênico, sendo este conceito, por isto, também conhecido como conceito de isolamento de espécie (CIE).

A Genética Marinha usa o CBE como base para utilização dos marcadores moleculares para estudos de sistemática e taxonomia. Para entender como isto é feito, basta lembrar o que foi visto no último parágrafo de **1.6.2. Estruturação**. Ou seja, se espécies são populações que não se intercruzam, então, indivíduos de duas espécies diferentes vivendo no mesmo local (simpatria) não estarão cruzando e, portanto, não estarão partilhando alelos que são exclusivos de cada uma destas espécies. Desta forma,

a amostragem da variação em alguns *loci* poderá evidenciar esta situação. Neste caso, os *loci* que mostrarem alelos exclusivos são chamados de *loci* diagnósticos.

Se o CBE define as espécies como grupos que não se intercruzam na natureza e são reprodutivamente isolados de outros grupos, ele independe da diferenciação morfológica entre os grupos. Desta forma, a aplicação da Genética Marinha a taxonomia e sistemática é especialmente importante nos casos em que existem dúvidas quanto ao *status* taxonômico de alguns tipos dentro de um grupo reconhecido como único, ou seja, para definição do *status* taxonômico dos chamados morfotipos (Knowlton 1993; 2000; Hillis et al. 1996; Klautau et al. 1999; Fukami et al. 2004; Silva 2009a; Chung & Chung 2012, Massana 2015) (ver Figura 1.31). Por exemplo, a espécie de anêmona *Anthopleura orientalis* apresentava dois morfotipos de cor em simpatria (morfótipo claro e morfótipo escuro) que, através do método de eletroforese de aloenzimas ficou evidente que eram, de fato, duas espécies diferentes (Manchenko et al. 2000). Dos 21 *loci* analisados, oito se mostraram diagnósticos!

A definição, pela Genética Marinha, de morfotipos como sendo espécies diferentes tem publicação certa, já que a classificação adequada das espécies é tarefa primária da biologia. Contudo, resultados negativos, ou seja, quando os morfotipos são mesmo morfotipos e não espécies diferentes, não têm publicação tão certa assim. Um exemplo que foge a esta regra é o estudo realizado com o ouriço-do-mar *Paracentrotus gaimardi*. Este ouriço-do-mar apresenta cinco morfotipos na costa brasileira que são caracterizados pelas cores dos seus espinhos (castanho, cinza, preto, rosa e verde). No sentido de investigar o *status* taxonômico destes morfotipos, marcadores moleculares de aloenzimas foram usados (Duarte et al. 2016). Contudo, neste caso, ao invés de terem sido amostrados apenas os morfotipos em simpatria, os autores resolveram amostrar os morfotipos também em alopatria. Desta forma, independentemente dos resultados em

termos taxonômicos (positivos ou negativos), os autores teriam resultados em relação à distribuição da variação gênica intra e interpopulacional. De fato, não foi possível demonstrar que os morfotipos de cor do ouriço-do-mar eram espécies diferentes, mas foram produzidos resultados interessantes quanto à distribuição da variação gênica (ver Quadro 1.29).

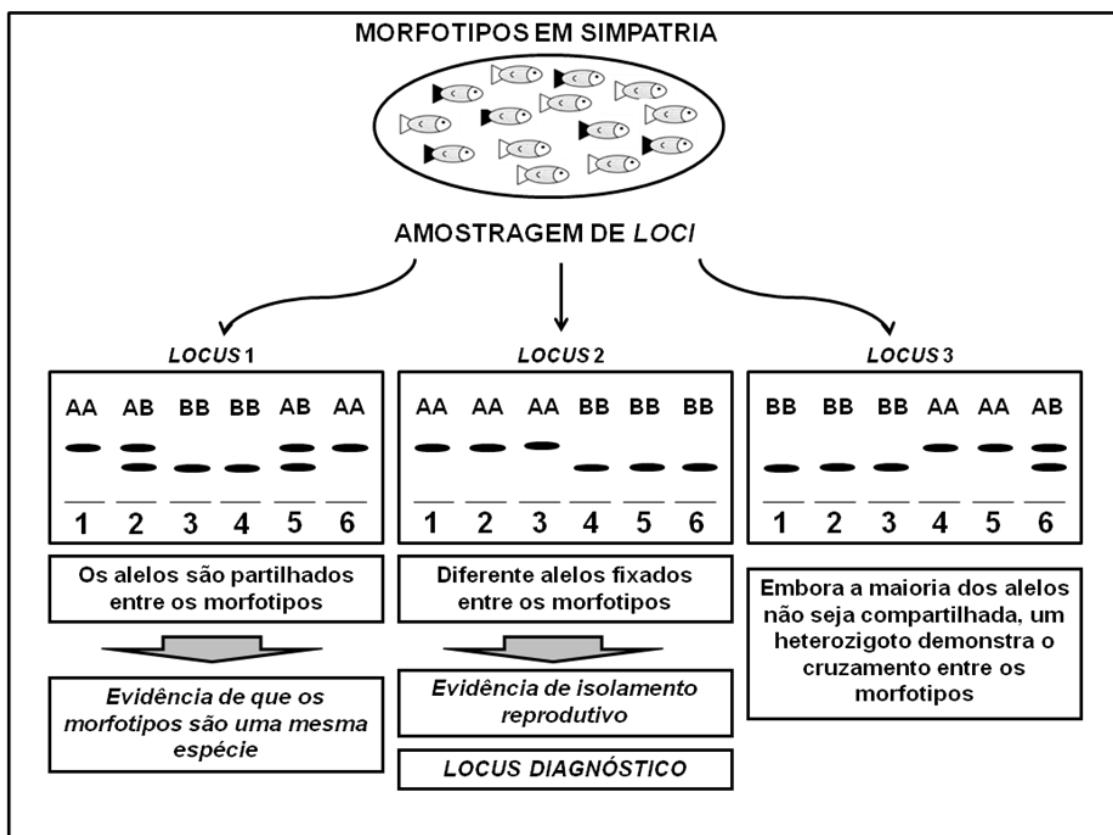
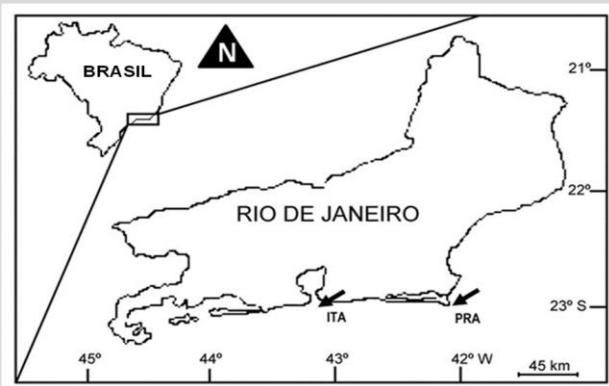


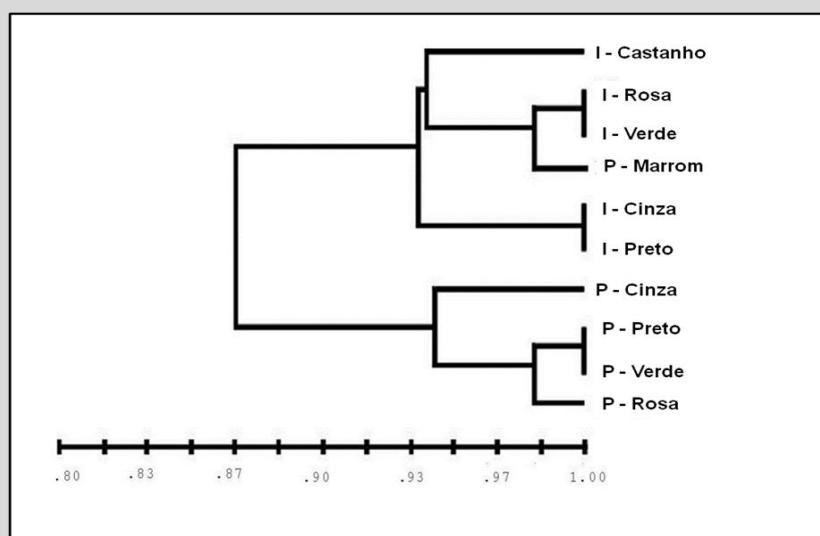
Figura 1.31. Representação de uma população composta por dois morfotipos de peixes (cauda branca e cauda preta) vivendo em sympatria. Estão exemplificados três *loci* usados para definir o status taxonomico dos morfotipos. O primeiro e o último *locus* não foram diagnósticos, uma vez que estão mostrando alelos partilhados pelos dois morfotipos, o que é uma indicação de ausência de isolamento reprodutivo entre eles. Contudo, o *locus* 2 é diagnóstico, já que indica que os morfotipos, mesmo vivendo em sympatria, apresentam isolamento reprodutivo, uma vez que não partilham alelos (nem um heterozigoto é observado).

Quadro 1.29. Estudo de sistemática bioquímica com a espécie de ouriço-do-mar *Paracentrotus gaimardi* no estado do Rio de Janeiro.

Foram amostrados indivíduos da espécie *Paracentrotus gaimardi* de ouriço do mar em duas localidades do estado do Rio de Janeiro (Prainha-Arraial do Cabo e Itaipu-Niterói) mostradas no mapa. O objetivo era verificar se os cinco morfotipos de cor (espinhos pretos, castanhos, cinzas, verdes e rosas) eram espécies diferentes.



Os resultados não evidenciaram nenhum *locus* diagnóstico, portanto, não foi possível definir os morfotipos como espécies diferentes. Contudo, como os morfotipos foram amostrados tanto em simpatria quanto em alopatria foi possível evidenciar que a variação gênica estava dividida entre os locais e não entre os morfotipos. O dendrograma de UPGMA abaixo ilustra bem este resultado ao mostrar um agrupamento pelos locais de amostragem e não pelos morfotipos (I = Itaipu, P = Prainha).



Duas coisas não foram ditas, ainda, sobre a utilização da Genética Marinha em taxonomia e sistemática. Primeiro é que ela oferece garantias sobre as diferenças, mas não é capaz de fazer o mesmo com as identidades. Ou seja, se um *locus* diagnóstico é encontrado, não existe dúvida de que os morfotipos investigados são espécies

diferentes. Contudo, se o *locus* diagnóstico não é encontrado, isto não significa que os morfotipos não são espécies diferentes, uma vez que um próximo *locus* amostrado pode ser aquele que se mostre diagnóstico. O que está se dizendo aqui é que a ausência de evidência não evidência a ausência.

A segunda coisa a ser dita é que a definição segura do *status* taxonômico de morfotipos só pode ser feita quando eles se encontram em simpatria. Quando os morfotipos se encontram em alopatria, os *locus* diagnósticos não podem ser considerados diagnósticos, uma vez que eles podem estar, apenas, mostrando a diferenciação geográfica (mais uma vez ver o último parágrafo de **1.6.2. Estruturação**). Ou seja, só existem *locus* diagnósticos em simpatria. Como se resolvem, então, problemas taxonômicos com alopatria? John Thorpe tentou dar uma resposta da Genética Marinha para definição do *status* taxonômico de morfotipos que não se encontram em simpatria. Ele analisou diversos trabalhos que estudavam, com aloenzimas, a estruturação genética entre populações da mesma espécie e as identidades genéticas entre espécies dentro do mesmo gênero, dentro da mesma família etc. Com base no conjunto de dados compilados, ele foi capaz de observar que populações coespecíficas (da mesma espécie) apresentavam identidades genéticas (IG) que, geralmente, se encontram acima de 0,80 e espécies congêneres (do mesmo gênero) apresentavam valores de IG que variavam entre 0,15 e 0,85. Estes valores não eram definidos com base em um conceito, eram, apenas, determinações empíricas, portanto, não são seguras, mas são boas indicações para se definir o *status* taxonômico de morfotipos que vivem em alopatria (Thorpe 1983; Thorpe & Solé-Cava 1994). A figura 1.32 apresenta os resultados de Thorpe.

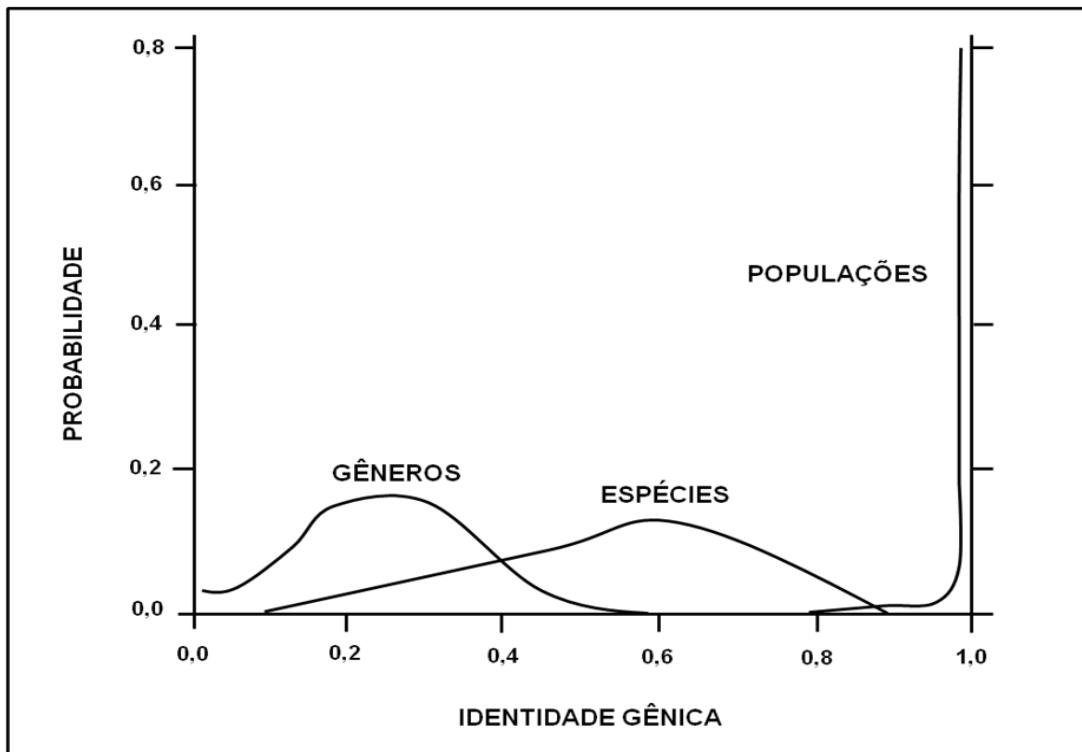


Figura 1.32. Gráfico de probabilidade da distribuição da Identidade Gênica de Nei para populações coespecíficas, congêneres e gêneros de uma mesma família.

Além das limitações do CBE para os casos de alopatria, ele apresenta outras e, por isso, tem sofrido severas críticas ao longo dos anos. A mais antiga crítica diz respeito a sua operacionalidade. Embora o CBE defina espécies pelo seu isolamento reprodutivo, a capacidade de intercruzamento da grande maioria das espécies é desconhecida. Por isto mesmo, na prática, a taxonomia baseada no CBE continua sendo feita com base na morfologia, ou seja, utilizando um conceito morfológico de espécie (CME). Do mesmo modo, o isolamento reprodutivo não é um critério universal, já que não se aplica para classificação de espécies com reprodução assexuada ou fósseis. Pelo CBE, espécies, por definição, não dizem. Contudo, um grande número de “boas” espécies (aqueles facilmente identificáveis e aceitas pelos sistemas como tal) intercruzam-se na natureza, em geral, e no mar (Gardner 1997). Por exemplo, as espécies de caranguejo *Menippe mercenaria* e *M. adina* na Flórida; as espécies de

bivalves *Mercenaria mercenaria* e *M. campechiensis* na Carolina do Sul e as algas *Fucus spiralis* e *F. vesiculosus* na Inglaterra. Por último, mas não somente, o CBE assume uma correspondência entre grupos reprodutivos e unidades morfológicas e ecológicas discretas (Hillis 1987). Contudo, isto não é verdade para muitos casos, especialmente no mar onde, com o advento da biologia molecular, um grande número de espécies crípticas (morfologicamente muito similares embora geneticamente distintas) tem sido revelado.

Um exemplo disto que se está falando é a espécie de esponja-do-mar *Cliona celata* (Xavier et al. 2010). *Cliona* é uma espécie ecologicamente importante já que é escavadora e desempenha papel na dinâmica dos ambientes costeiros. Ela era considerada uma espécie cosmopolita, com presença extensivamente documentada nos oceanos Atlântico, Índico e Pacífico, bem como no Mediterrâneo e outros mares fechados. Contudo, quando o *status taxonomico* de diferentes formas de crescimento foi investigado em populações simpátricas e alopátricas do nordeste do Atlântico e do Mediterrâneo ficou evidente que se tratava de pelo menos quatro espécies diferentes. A análise foi feita com marcadores de mtDNA (COI e Atp8) e nDNA (28S). *C. celata* passou a ser reconhecida, então, como um complexo de espécies para o qual as diferentes espécies são, praticamente, indistinguíveis. Isto porque, taxonomicamente, esponjas são classificadas com base nas suas espículas e, as espécies do complexo, se distinguem pelo modo de crescimento. Assim, *Cliona celata* que era cosmopolita, passou a ter uma distribuição mais restrita: duas espécies vivendo na costa européia do Oceano Atlântico e outras duas vivendo no Mediterrâneo.

A técnica de PCR e a sua capacidade de amplificar segmentos de DNA que são espécie-específicos têm um impacto importante também para a atividade de identificação de ovos, larvas e jovens no plâncton, que é um trabalho laborioso e muitas

vezes impossível. Esta abordagem já foi bem sucedida em estudos de identificação de larvas e jovens de mexilhões, ictioplâncton e bactérias de sedimento, entre outros. Do mesmo modo, a capacidade da técnica em discriminar as espécies através de marcadores moleculares, tem sido útil, também, no controle e fiscalização da pesca. A identificação da espécie de origem da carne apreendida em barcos pesqueiros já foi feita para fiscalização da pesca ilegal de baleias e moluscos, no que tem sido chamado de genética forense da pesca. A genética marinha tem sido importante, também, para estabelecer ou auxiliar o entendimento de relações evolutivas entre os mais diversos grupos de organismos marinhos, de diatomáceas a mamíferos marinhos.

Dito isto, a taxonomia e sistemática feita com base na Genética Marinha não substituiu e nem irá substituir a taxonomia e sistemática clássica. A taxonomia e sistemática molecular é uma ferramenta que permite a identificação de espécies nos casos em que as características utilizadas pela sistemática clássica não são suficientes, como na distinção de espécies morfológicamente simples. Contudo, mesmo aí, técnicas de morfometria e citogenética, que são mais baratas, podem, algumas vezes, fazer o serviço. Por exemplo, no leste da Australia foram coletados 97 espécimes dos platelmintos *Gyratrix hermaphroditus* para análises morfológicas e citogenéticas. Os resultados evidenciaram a existência de, pelo menos, oito espécies crípticas todas caracterizadas por diferentes números e tamanhos de cromossomos, bem como diferentes índices centrométricos (definidos em função da posição do centrômero) (Curini-Galletti & Puccinelli 1998). As diferentes espécies apresentaram, também, formas diferentes para o órgão reprodutor masculino. As diferenças foram observadas tanto em simpatria quanto em alopatria, os grupos em simpatria consitiam claramente em espécies diferentes. No caso da alopatria, contudo, a caracterização das espécies ainda é problemática.

1.6.4. Genética Ecológica

A genética ecológica tenta demonstrar como a seleção natural atua nas populações de organismos marinhos. Esta é uma tarefa que, muitas vezes, demanda a interação da Genética Marinha com outras disciplinas como, por exemplo, a ecologia, a fisiologia e a bioquímica. Pois bem, um fato já estabelecido em relação às populações naturais é que os seus níveis de variação gênica, em nível molecular, são altos. Mais que isso, a explicação básica para este fato é a hipótese neutralista (ver *1.2.1.3. Controvérsias evolutivas*). Por outro lado, os selecionistas buscam explicar a variação gênica a partir da sua relação com a variação ambiental. Desta forma, em muitos aspectos, as explicações da genética ecológica e do selecionismo se sobrepõem; entretanto, genética ecológica e selecionismo não são a mesma coisa! A genética ecológica surgiu antes da teoria selecionista, já na década de 1920, e era a denominação dada por Edmund B. Ford aos seus estudos sobre a ação da seleção natural em populações naturais. Os dois exemplos clássicos desse campo de estudo são aqueles referentes à mariposa *Biston betularia* e o melanismo industrial e sobre o padrão de listras no caracol *Cepaea nemoralis*, nenhum dos quais é marinho, contudo, os primeiros estudos em Genética Marinha se devem, também, a Ford (ver *4.2.2. O Começo*). A teoria selecionista data da década de 1960, com a entrada em cena dos marcadores moleculares (ver *1.2.1.3. Controvérsias evolutivas*).

Se os estudos sobre aquilo que se compreende como genética ecológica é mais antigo, o termo só surgiu na década de 1950, mais exatamente em 1956, quando o *Ecological Genetics Group* (Grupo de Genética Ecológica) realizou um encontro em Aberystwyth, no País de Gales. Contudo, é creditada a Ford a criação do termo, a partir da publicação do seu livro *Ecological Genetics*, em 1964. Mais que isso é de Ford,

também, uma primeira definição formal da área, como aquela que estuda as adaptações das populações naturais ao seu ambiente. Embora muito clara esta definição esconde dificuldades que advêm da definição de adaptação (Gould & Lewontin 1975) . Esta palavra possui muitas conotações diferentes, tanto no senso comum, quanto em Ciência. Por exemplo, diz-se que pessoas estão bem adaptadas a algum local ou situação quando elas estão confortáveis ou desenvolvem laços sociais. Nesse caso, adaptação é quase um sinônimo de bem-estar, satisfação ou, mesmo, felicidade. De modo semelhante, em disciplinas como a fisiologia, por exemplo, a palavra pode ser usada como sinônimo de ajustamento fenotípico de um organismo individual ao seu ambiente, ou seja, um processo que seria mais bem denominado de aclimatação fisiológica (Ford 1975).

Em termos da genética ecológica, duas coisas são fundamentais para que se defina adaptação. Primeiro, é necessário que exista variação gênica, que é o material sobre o qual a seleção natural irá agir e, segundo, a seleção natural. Portanto, adaptação diz respeito, sempre, a seleção natural e a variação gênica. Isso tem consequências muito importantes para a compreensão do fenômeno. Por exemplo, observando um caráter que é utilizado para uma função específica por um determinado organismo, é possível imaginar soluções mais adequadas do que aquelas que se vê. Contudo, as adaptações são, sempre, produto da história evolutiva das espécies e, como tal, elas são limitadas pela variação gênica presente nas populações. A seleção natural não inventa ou determina a variação necessária para dada circunstância, ela simplesmente aproveita o material disponível. A figura 1.33 resume o que foi dito até aqui, ou seja, o enquadramento teórico e prático no qual os estudos de genética ecológica se dão.

A utilização do argumento de que a seleção natural é a força determinante das alterações nas frequências gênicas em um ou mais loci, depende da observação de alguns fatos: (1) variação nas frequências gênicas ao longo do tempo que não possa ser

explicada apenas pela ação do acaso; (2) um desequilíbrio de ligação entre *loci* que possa ter sido gerado por seleção natural; (3) a correlação da variação de frequência com algum fator ambiental e (4) que o padrão se repita em outras unidades evolutivas. Mesmo assim, a aceitação de que um polimorfismo molecular é mantido por seleção natural depende, ainda, que seja demonstrado que: (a) os alelos em questão produzem fenótipos diferentes; (b) que estas diferenças fazem sentido no ambiente e (c) que estas diferenças tenham algum tipo de impacto na sobrevivência dos indivíduos. Existem, pelo menos, três bons exemplos do desenvolvimento deste programa de pesquisa em organismos marinhos. Um deles é o do *locus Lap*, em moluscos, que já foi discutido aqui (ver **1.6.1. Variação Gênica** e **1.6.2. Estruturação**).

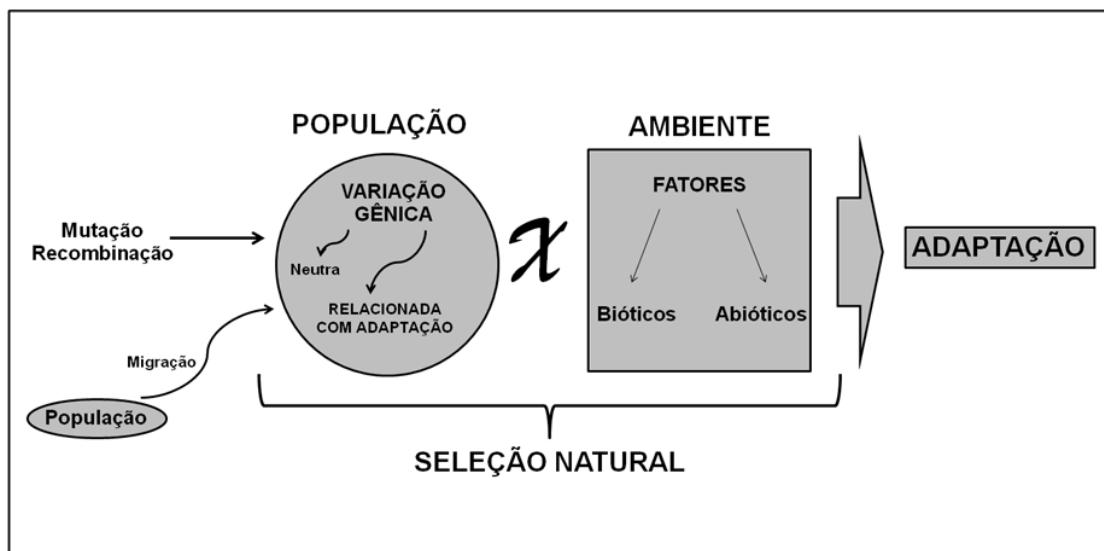


Figura 1.33. O enquadramento teórico da genética ecológica. São mostradas as fontes de variação gênica que são intrínsecas (mutação e recombinação) e extrínsecas (migração). Contudo, desta variação, aquela que interessa aos estudos de genética ecológica são aquelas que são sensíveis ao ambiente, ou seja, estão relacionadas com a adaptação. A seleção (natural) dos variantes populacionais herdáveis (variação gênica) é que vai determinar a adaptação.

O outro exemplo clássico diz respeito aos efeitos do polimorfismo da enzima Transaminase Glutamato Piruvato (GTP) na regulação do volume celular no copépodo *Tigriopus californicus* (Burton & Feldman 1983). Da mesma forma que no caso da LAP

em moluscos, GTP tem um papel relevante na osmorregulação, contudo, neste caso, a função enzimática esta relacionada à resposta dos copépodos ao estresse hiperosmótico. Existem dois alelos mais comuns do *locus GTP* nas populações de *Tigriopus*, GTP^F (F de *fast*, em inglês, ou seja, rápido, em português) e GTP^S (S de *slow*, em inglês, ou seja, lento, em português). O alelo GTP^F determina a produção da forma aloenzímica que tem uma maior atividade que aquela produzida pelo alelo GTP^S . Desta forma, indivíduos com genótipo GTP^S/GTP^S estão sujeitos a uma maior mortalidade das suas larvas em condições de estresse hiperosmótico.

O terceiro exemplo clássico é do peixe teleosteio *Fundulus heteroclitus* (Powers & Place 1978; Place & Powers 1984; Gonzalez-Vilasenor & Powers 1990; Ropson et al. 1990). Um grande número de evidências indica que o polimorfismo da enzima Lactato Desidrogenase (LDH) neste peixe está envolvido nos mecanismos de adaptação a diferentes temperaturas no ambiente marinho. Os dois alelos mais comuns do *locus LDH-B* ($Ldh-B^a$ e $Ldh-B^b$) apresentam uma diferença de frequência que é estatisticamente significativa, estável no tempo e que acompanha a linha norte/sul ao longo da costa atlântica dos Estados Unidos. As diferenças podem ser explicadas pelas atividades cinéticas das enzimas produzidas pelos alelos alternativos do *locus*: O alelo $Ldh-B^b$, para o qual a frequência é maior nas águas frias do norte, determina a produção da forma aloenzimática que tem maior atividade em temperaturas mais baixas (altas latitudes). O alelo alternativo ($Ldh-B^a$), por outro lado, apresenta maior frequência nas águas quentes do sul e a sua forma aloenzimática é mais eficiente em temperaturas mais elevadas (baixas latitudes). Assim, o polimorfismo e as diferenças de frequência observadas são mais bem explicados por seleção natural do que por qualquer outra das três (mutação, deriva e migração) forças evolutivas.

Embora estudos bioquímicos, morfológicos e fisiológicos realizados com os três exemplos clássicos da genética ecológica tenham evidenciado clines consistentes de frequências gênicas (de aloenzimas) em relação a gradientes ambientais, clines não são, necessariamente, produzidos por seleção natural. Algumas vezes, a história evolutiva das espécies é a chave de interpretação de um cline. No caso da espécie *Fundulus heteroclitus*, por exemplo, foram realizados estudos tanto com mtDNA quanto com microsatélites (Adams et al. 2006) e os resultados evidenciaram que mesmo para estes marcadores, considerados neutros, eram encontrados padrões correspondentes aqueles vistos para os *loci* de aloenzimas.

Os estudos utilizando mtDNA revelaram dois grupos de haplótipos, um dos quais, mais homogêneo, correspondendo, geograficamente, às populações do norte. O outro grupo, mais heterogêneo, correspondia às populações do sul. No caso dos microsatélites, foram utilizados oito *loci*, todos mostraram o mesmo cline de frequência norte-sul, bem como um aumento na diversidade de alelos nas populações do sul. As diferenças entre as regiões norte e sul coincidem com uma barreira zoogeográfica que limitou ou impediu o fluxo gênico da espécie *Fundulus heteroclitus* durante o último período de glaciação. Nestes casos, portanto, a explicação para os clines observados reside no modelo de isolamento por distância. Ou seja, é um componente histórico e demográfico que está produzindo o padrão de variação observado.

A espécie *F. heteroclitus* é um exemplo para o qual foi possível compor dados de diferentes marcadores moleculares para entender a origem dos clines observados, tanto aqueles presumivelmente neutros (mtDNA e microsatélites) quanto aqueles para os quais existem boas evidências de seleção natural (*Ldh* em aloenzimas) (Schmidt et al. 2008).

Com a disponibilidade atual de uma ampla gama de marcadores moleculares de DNA, outra abordagem da genética ecológica tem sido possível, os chamados *scans* genômicos. Neste caso, um grande número de *loci* pode ser amostrado em extremos de um gradiente ambiental e os F_{ST} calculados. Pela hipótese neutralista, espera-se uma homogeneidade dos valores de F_{ST} . Se alguns marcadores apresentarem valores de F_{ST} excepcionalmente elevados, isto é uma indicação de que a seleção natural pode estar envolvida neste padrão. Certamente, não diretamente sobre os marcadores de DNA que são, supostamente, neutros. Mas, provavelmente, em outros *loci* que se encontram ligados a eles. Aqui o termo “marcador molecular” fica bem claro: os *loci* de DNA estão marcando que alguma coisa anômala está ocorrendo em função do fator ambiental. Embora, nestes casos, não seja possível identificar os mecanismos causais que estão determinando o fenômeno, é interessante demonstrar que a seleção natural está deixando as suas marcas genéticas. Esta abordagem foi utilizada com o molusco gastrópodo *Littorina saxatilis* comparando médio e sublitoral com um *scan* de 306 marcadores de AFLP (McInerney et al. 2009). Foi possível encontrar muitos dos marcadores para os quais os valores de F_{ST} eram maiores do que aquilo que se esperaria pela hipótese neutralista. A seleção natural é a explicação, embora não direta, para o padrão encontrado.

Uma variação ambiental dramática a qual os organismos marinhos têm estado sujeitos, especialmente nos últimos séculos, é a poluição antropogênica. A seguir será visto como a Genética Marinha tem colaborado, senão para resolver, para evidenciar os problemas da poluição no ambiente marinho.

1.6.5. Genética e Poluição

Os ecossistemas marinhos costeiros vêm sofrendo um intenso processo de degradação em todo o mundo. Relatos de grandes contaminações, como os casos de derramamento de petróleo, são amplamente divulgados pela mídia. Nestas ocasiões, a fonte poluidora é conhecida e as consequências são visíveis, como mudanças abruptas na paisagem e a morte massiva de animais. Entretanto, o tipo mais comum de poluição consiste em descargas constantes de efluentes de origem industrial e doméstica que comprometem os ambientes de forma menos evidente, mas nem por isso, menos tóxica. Os poluentes podem provocar diversas alterações nos organismos, incluindo alterações genéticas (Dixon & Wilson 2000).

A poluição marinha é capaz de causar dois tipos de efeito na constituição genética dos organismos. Primeiro, a ação dos poluentes pode causar danos ao material hereditário, trazendo como consequência mutações genéticas pontuais (um ou poucos nucleotídeos) ou aberrações cromossômicas. O resultado destes eventos pode ser a morte do indivíduo, o desenvolvimento de neoplasias ou de mosaicos de células no organismo. Em qualquer dos casos ocorre uma diminuição da adaptabilidade dos indivíduos afetados que, dependendo da extensão do dano genético, pode acarretar uma ameaça de extinção da população ou mesmo da espécie como um todo.

Outro efeito que os poluentes podem ter na constituição genética das populações advém do fato de que, como eles alteram o ambiente marinho, podem acabar agindo como agentes seletivos. Nestes casos, a população sofre um ajuste às novas condições ambientais, o que pode significar a modificação das frequências gênicas em *loci* específicos ou a alteração da constituição genômica como um todo.

Em Genética Marinha, os efeitos dos poluentes nos organismos marinhos podem ser estudados, por exemplo, através da citogenética. Danos no material hereditário

podem ser mensurados através de marcadores como os *micronucleus* (micronúcleos em Português). Os micronúcleos são estruturas produzidas nas células por ação, geralmente, de compostos clastogênicos (substâncias ou processos que causam rupturas cromossômicas) ou aneugênicos (substâncias ou processos que afetam a divisão celular). Mexilhões do gênero *Mytilus* spp. foram utilizados para avaliar o nível de poluição ambiental no Mar Báltico, em localidades que diferem tanto quanto aos níveis de poluição ambiental quanto em termos de condições hidrográficas (salinidade, oxigênio, temperatura) (Barsiene et al. 2006). Os resultados da comparação das localidades evidenciaram que as diferenças na ocorrência de micronúcleos estava relacionada com o grau de contaminação do habitat por Bifenilas policloradas ou PCBs (do inglês, *Polychlorinated biphenyls*) e Hidrocarbonetos Policíclicos Aromáticos ou PAHs (do inglês, *Polynuclear Aromatic Hydrocarbons*). A contagem de micronúcleos permite, portanto, uma avaliação da quantidade de transformações morfológicas do material genético. Outros marcadores, como as aloenzimas, avaliam mudanças nas frequências gênicas.

Trabalhos realizados com a craca *Balanus amphitrite*, demonstraram que as frequências de vários alelos da aloenzima *Pgi* estavão correlacionadas com os níveis de poluição. Do mesmo modo, no mexilhão *Perna viridis*, os metais pesados, cádmio e cobre, foram identificados como os principais responsáveis pelas alterações de frequências alélicas neste *loci*. Outros resultados interessantes a respeito da relação entre variação alélica de aloenzimas e poluição é a correlação que já foi observada entre heterozigosidade e poluição. Em Israel, espécies de gastrópodos dos gêneros *Monodonta*, *Littorina* e *Cerithium* apresentaram uma correlação positiva entre níveis de diversidade genotípica (heterozigosidade, grau de polimorfismo e número efetivo de alelos) e resistência a poluentes como metais pesados, óleos e detergentes.

Estudos experimentais também podem ser realizados e, com raras exceções, são desenvolvidos com modelos que diferem daqueles para os quais foram realizadas as observações diretas no campo, o que torna a comparação difícil. Do mesmo modo, as concentrações de poluentes utilizadas em laboratório são, geralmente, muito superiores àquelas que seriam realísticas no campo. Contudo, em alguns casos, boas correlações puderam ser obtidas entre resultados experimentais e observações no campo. Estudos com os gastrópodos *Littorina neritoides* e *L. punctata*, realizados ao longo da costa de Israel, demonstraram que alelos que eram negativamente selecionados sobre a ação do mercúrio no laboratório, eram os mesmos que estavam em frequência significativamente baixa em áreas poluídas quando comparados com áreas não poluídas. Outro bom exemplo de parceria entre estudos de campo e laboratório foi realizado com a craca *Balanus amphitrite* na Laguna de Veneza, onde os efeitos da poluição térmica e química foram estudados em três populações naturais por um ano. A composição específica de metais pesados no sedimento e nos tecidos das cracas foi identificada. Posteriormente, ensaios de laboratório determinaram que a temperaturas por volta de 20°C e sobre a ação de diferentes concentrações de cobre e zinco, a atividade da enzima Pgi era significativamente maior em homozigotos que em heterozigotos nos indivíduos das três populações.

Além dos marcadores que avaliam diferenças na forma do material genético e mudanças nas frequências gênicas causadas pela poluição, existem, também, aqueles que avaliam alterações nos níveis de transcrição de genes relacionados com a transformação de determinados poluentes (ver *Microarrays* em **1.3.7. Ómicas**). O citocromo P4501A é uma família de proteínas responsável pela transformação de compostos orgânicos, como PCBs e PAHs. No estado de Santa Catarina os níveis de transcrição de quatro isoformas de P4501A foram avaliados por meio de PCR

quantitativa de transcriptase reversa (qRT-PCR) com ostras da espécie *Crassostrea brasiliiana* transplantadas e mantidas durante 24 horas em três áreas potencialmente contaminadas por esgoto sanitário. Os resultados evidenciaram que os níveis de transcrição do gene CYP2AU1 (uma das isoforma de P4501A) aumentaram, evidenciando, desta forma, uma correlação positiva com a poluição (Bainy & Marques 2001; Pessattia et al. 2016).

Os estudos a respeito de genética e poluição não são muitos e os resultados muitas vezes são contraditórios. Estes dois problemas se devem, principalmente, ao fato da poluição marinha ser um evento contingencial, histórico e multifatorial. Por conta disto, efeitos observados em uma região não podem, geralmente, ser extrapolados para outras e o agente causal das alterações na constituição genética das populações é obscurecido pela grande quantidade de agentes envolvidos no fenômeno da poluição, ou na relação entre dois ou mais agentes (naturais e/ou poluentes) que podem atuar de modo simples ou sinergístico. A figura 1.34 resume a abordagem da Genética Marinha aplicada aos problemas de poluição do ambiente marinho.

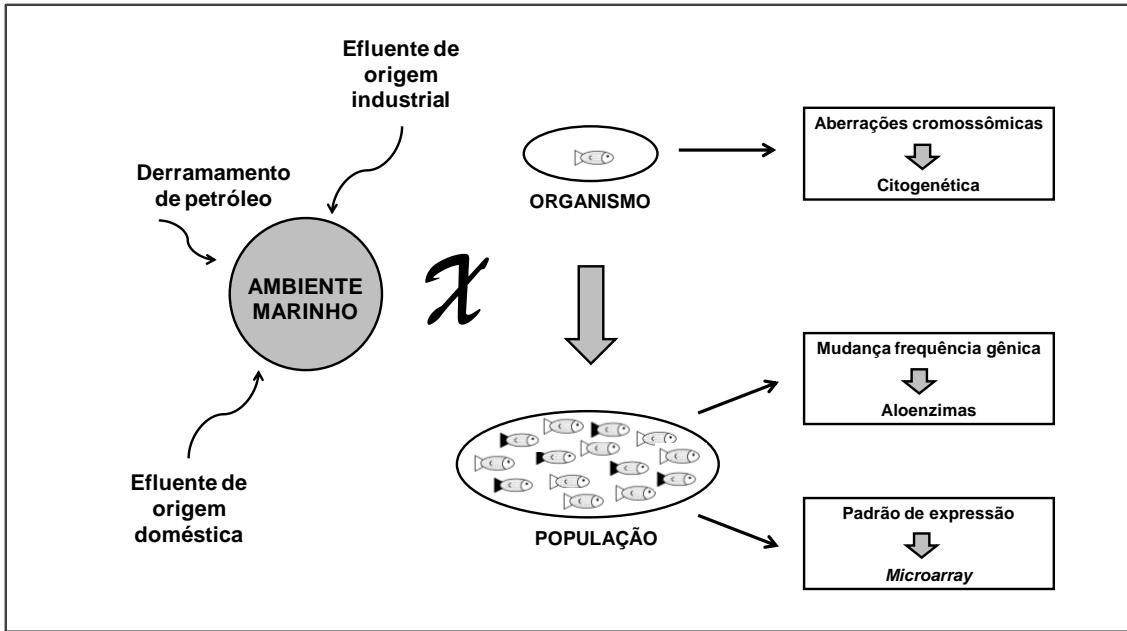


Figura 1.34. Esquema indicando algumas abordagens da Genética Marinha para os estudos sobre poluição do ambiente marinho.

1.6.6. Genética Pesqueira

A pesca é praticada, principalmente, como uma atividade extrativista que produz efeitos na diversidade e abundância das espécies marinhas, podendo levar a mudanças significativas na estrutura das comunidades locais. É neste sentido que a atividade pesqueira demanda políticas de manejo que preservem os recursos e garantam sua exploração em longo prazo (Hilsdorf 2006; EMBRAPA 2012). O manejo de um recurso pesqueiro depende de conhecimento a respeito de variáveis fisiológicas, comportamentais e ecológicas das espécies de interesse, de modo que se possa determinar como suas populações sobrevivem e se reproduzem nos diferentes ambientes e locais. Mais que isso, é fundamental se saber de quantas e quais espécies se está falando. Ou seja, é fundamental conhecer a estrutura genética populacional de uma dada espécie ou, como se diz em biologia pesqueira e, também, em Genética Marinha definir o estoque que sofre o esforço de pesca (Ryman & Utter 1987; Waples 1987; Ward et al. 1992; Ward et al. 1994; Ward & Grewe 1994; Carvalho & Hauser 1995; Waples et al. 2008). Além da identificação das unidades reprodutivas (estoques), a aplicação da

Genética Marinha a biologia da pesca envolve, ainda, a estimativa de tamanhos de população e a compreensão da relação entre variáveis como heterozigozidade e crescimento (ver **1.6.7. Aquicultura**). De fato, tudo isto está relacionado com uma coisa só: a capacidade da Genética Marinha de descrever a variação gênica e entender o seu padrão de distribuição e organização nas populações naturais (ver **1.4. Conceitual**).

Com relação aos procedimentos de delimitação de estoques, a abordagem mais usada é a amostragem extensiva na área de pesca, seguida de comparações das frequências gênicas em sistemas polimórficos entre os vários pontos de coleta, visando avaliar se todos eles fazem parte de apenas uma grande população panmítica (i.e., se existe apenas um estoque) ou se a variabilidade está distribuída de maneira hierárquica em vários estoques diferentes. No primeiro caso, o manejo do recurso deve ser feito de maneira unificada entre todos os grupos que o exploram. Isto porque a exploração de um recurso pesqueiro pode envolver desde diferentes frotas de barcos até diferentes nações. No segundo caso, o manejo deve ser feito separadamente para cada unidade de pesca (ou subpopulação). Uma vez que, se a espécie se apresenta estruturada, o impacto da exploração será sentido em cada unidade na dependência do seu tamanho efetivo. Ou seja, vai depender muito dos níveis de migração e mistura (os níveis de fluxo gênico) entre as subpopulações.

A Genética Marinha tem obtido relativo sucesso no estudo de populações e na delimitação de estoques de várias espécies de peixes marinhos comerciais (por exemplo, para as espécies *Thunnus thynnus*-Atum azul, *T. alalunga*-Atum branco e *Sarda sarda*-Bonito) utilizando diferentes marcadores moleculares. Estudos de genética pesqueira (neste capítulo, a Genética Marinha aplicada a pesca) foram fundamentais, por exemplo, para resolver a disputa entre os governos da Austrália e da Tailândia em relação ao manejo dos estoques de cação pelos dois países e, também, entre os governos dos EUA

e do Canadá sobre a pesca da perca. Desta forma, estudos de delimitação de estoques têm, muitas vezes, importância, também, para assuntos geopolíticos, como a atribuição de responsabilidades na administração dos recursos naturais. Entretanto, na prática, as decisões não são tomadas por geneticistas em solo, mas por grupos de cientistas, políticos e administradores que precisam trabalhar juntos para definir melhor as estratégias que servirão de base para elaboração de leis com os objetivos de manejo e conservação dos recursos pesqueiros.

No Brasil, a Genética Marinha têm sido importante, desde a década de 1970, na definição do número de estoques pesqueiros da espécie *Micropogonias furnieri*, a corvina. O primeiro trabalho preocupado com o problema desta espécie foi realizado com biometria e morfometria e definiu, então, a existência de dois estoques de corvina para costa sul-sudeste do Brasil. Trabalhos posteriores, já usando marcadores moleculares (aloenzimas e mtDNA) não foram capazes de definir a existência de mais de um estoque. Contudo, mais recentemente, com a utilização de marcadores moleculares mais variáveis (microssatélites e EPIC) e amostragem mais extensiva (coletas realizadas entre o Rio de Janeiro e o Rio Grande do Sul, mas, também, uma população do estado do Pará) foi possível demonstrar a existência de três estoques de corvina muito bem diferenciados: norte do Brasil (Pará), regiões sudeste e inicio da região sul e outro mais ao sul. Estes resultados definiram que as pescarias desta espécie devem ser geridas com ações específicas para cada um dos três estoques (Vasconcellos et al. 2015).

A existência de um estoque de *M. furnieri* na costa norte brasileira era esperado devido à grande distância geográfica da população do norte e das populações do sudeste do Brasil (mais de 5000 km). Contudo, mais que isso, padrões semelhantes têm sido encontrados para outras espécies sugerindo que o fluxo gênico entre as populações de peixes do norte e do sudeste do Brasil pode ser muito restrito. Um fator que pode

explicar a diferenciação entre essas duas regiões é a prevalência de correntes oceânicas ao longo da costa atlântica da América do Sul. Por outro lado, a presença de dois estoques distintos entre as regiões sudeste e sul do Brasil pode estar relacionada com a influência da corrente quente do Brasil e da corrente fria das Malvinas, que formam a convergência subtropical quando se encontram no sul do Brasil. Biólogos marinhos já haviam apontado a existência desses dois estoques que só foram reconhecidos pela Genética Marinha recentemente. Este fato chama atenção para o fato de que a conservação e o manejo dependem de um bom diálogo entre os diferentes campos de conhecimento interessados nestes problemas.

Embora muitas vezes os marcadores moleculares sejam a escolha natural para resolver problemas de delimitação de estoques, isto nem sempre é necessário. As vezes apenas com marcadores de forma, como as técnicas de morfometria são suficientes para resolver o problema. Por exemplo, estoques do carapau do mediterrâneo, *Trachurus mediterraneus*, foram estudados nos mares Negro, Marmara, Egeu e Mediterrâneo (Turan 2004). Os resultados sugeriram que as populações poderiam estar estruturadas e, mais que isso, apresentando um recrutamento local. Estes resultados foram suficientes para que houvesse a indicação de uma gestão independente para as diferentes áreas, uma vez que um esgotamento de qualquer dos estoques não poderia ser compensado pela imigração, ao menos em velocidade rápida o suficientemente para evitar a extinção. Contudo, algumas vezes, para uma identificação confiável de um estoque é necessária uma variedade de métodos, já que resultados contraditórios podem ser encontrados para um par de métodos ou entre os resultados de um método e dados biológicos, culturais etc. Na figura 1.35 é apresentado um esquema geral daquilo que foi apresentado aqui sobre a aplicação da Genética Marinha a pesca.

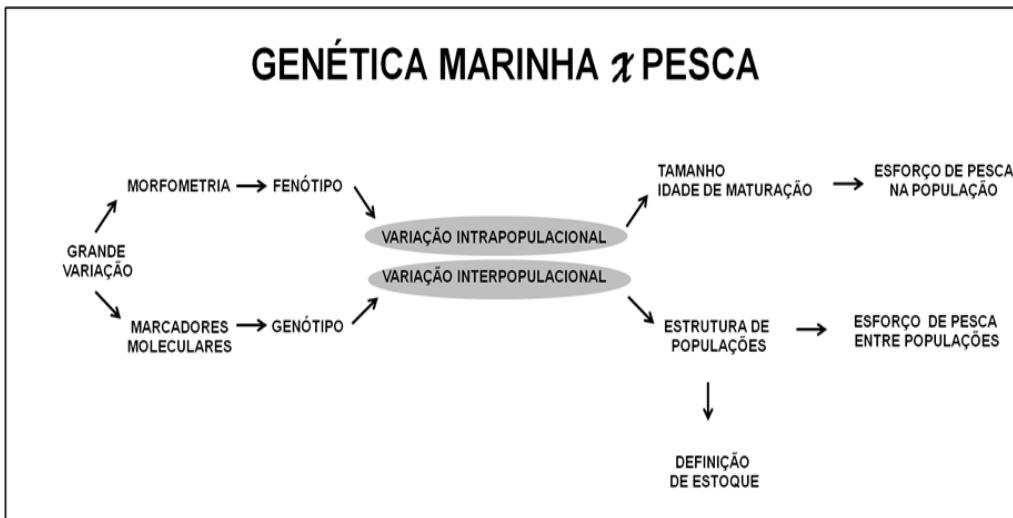


Figura 1.35. Esquema indicando a abordagem da Genética Marinha aplicada à pesca. Detalhes de estudos nessa perspectiva estão descritos no texto.

No Brasil, outro recurso pesqueiro importante é a espécie de mexilhão *Perna perna* que ocorre do Espírito Santo ao Rio Grande do Sul, fazendo parte do cardápio de restaurantes através de todo país (o mexilhão é exportado para as áreas onde não ocorre) (Kant de Lima & Pereira 1998). Além do que, sua exploração direta de bancos naturais constitui um importante recurso alimentar e econômico para populações de baixa renda, os marisqueiros, que sobrevivem do seu consumo e comercialização. As atividades de cultivo de *Perna perna* ainda são incipientes no Brasil, contudo, se desenvolvem rapidamente como atestam as atividades de cultivo no Rio de Janeiro, São Paulo e Santa Catarina. Em função do desenvolvimento das atividades de cultivo do mexilhão, a contradição de interesses começa a se instaurar entre a tradicional exploração de bancos naturais e a moderna atividade de cultivo, principalmente no que diz respeito à obtenção de sementes. Este é um caso para o qual a genética pesqueira tem interseção com a aquicultura, para qual as contribuições da Genética Marinha são expostas a seguir.

1.6.7. Aquicultura

A aquicultura possui um importante papel na produção de recursos alimentares que em outro momento eram obtidos através da exploração das populações naturais (Oliveira 2009). Esta atividade refere-se ao cultivo de diversos organismos aquáticos como plantas, moluscos, crustáceos e peixes, sendo que a intervenção ou manejo do processo de criação é imprescindível para o aumento da produção. A herança de traços desejáveis na aquicultura tem sido a base para a criação, com programas concentrados em selecionar os reprodutores com as características mais convenientes. Estes programas levaram a sistemas de aquicultura eficientes, produtivos e rentáveis. Neste caso, a Genética Marinha tem desempenhado papel significativo no desenvolvimento de estratégias que envolvem, principalmente e não obrigatoriamente isoladas, o melhoramento genético, a seleção de matrizes e a manutenção da qualidade dos cultivos.

A base das estratégias de melhoramento genético de organismos marinhos é mais ou menos semelhante àquelas utilizadas para os organismos terrestres. A diferença, neste caso, reside no fato de que os organismos marinhos têm níveis mais elevados de variação gênica e, portanto, oferecem mais material para programas de seleção artificial. A seleção de matrizes é um processo em que os indivíduos com fenótipos particulares são identificados e usados como reprodutores para produzir descendentes com a característica desejada. Os programas de seleção incluem a seleção em massa e a seleção de famílias. Na seleção em massa, o desempenho de todos os indivíduos é comparado e a seleção é baseada no desempenho de cada um, desconsiderando a filiação. Na seleção de famílias, o desempenho médio das famílias é comparado e famílias inteiras são selecionadas. Para garantir a eficácia de um programa de seleção, a herdabilidade da característica em foco deve ser tão grande quanto possível.

Muitas características podem ser correlacionadas, positiva ou negativamente. Por exemplo, a taxa de crescimento rápido é frequentemente correlacionada com uma taxa de conversão alimentar mais eficiente em algumas espécies. Desta forma, a seleção para organismos de crescimento rápido também seleciona para organismos com uma melhor taxa de conversão alimentar. No entanto, se os traços estão correlacionados negativamente, a seleção para uma característica pode afetar negativamente o outro. Por exemplo, o crescimento rápido pode ser negativamente correlacionado com a capacidade de reprodução, de modo que a seleção para os organismos de crescimento rápido poderia reduzir a capacidade reprodutiva. Portanto, os programas de seleção devem ser cuidadosamente monitorados para determinar se há alguma correlação entre as características e, portanto, evitar resultados indesejáveis.

Se a herdabilidade da característica desejada é muito baixa, os métodos de seleção de caracteres fenotípicos não serão uma boa opção. Neste caso, programas de seleção artificial dependem de que variação gênica seja grande, de modo a propiciar combinações variadas de alelos. Uma estratégia para aumentar a variação gênica e, portanto, o número de combinações possíveis, é o cruzamento de organismos com diferentes histórias genéticas, um processo chamado de hibridização. A hibridização pode ser entre indivíduos de diferentes cepas, linhagens ou populações. Neste caso, ela é chamada de hibridização intraespecífica, uma vez que os pais são da mesma espécie. Alternativamente, os pais podem ser de diferentes espécies. Neste caso, a hibridização é interespecífica. A estratégia de hibridização é empírica, ou seja, é feita através de tentativa e erro, já que os híbridos podem, ou não, ter características desejáveis. Contudo, híbridos apresentam, obrigatoriamente, mais variação gênica que os pais e isso, por si só, é uma coisa desejável em programas de aquicultura, inclusive por que alguns estudos têm demonstrado uma correlação entre heterozigosidade (uma das

medidas de variação gênica, ver **1.4.1. Variação Intrapopulacional**) e características importantes para aquicultura de organismos marinhos.

A heterozigosidade é estimada em organismos de cultivo usando-se marcadores moleculares (aloenzimas, mtDNA, nDNA). No caso das estimativas de heterozigosidade baseadas em aloenzimas, existem dois grupos de hipóteses para explicar a correlação entre heterozigosidade e melhor desempenho dos organismos. No primeiro caso, é apontado um envolvimento direto das aloenzimas no fenômeno. Assim, a correlação positiva entre heterozigosidade e componentes da adaptação estaria relacionada à maior eficiência média do metabolismo basal nos indivíduos mais heterozigotos (Garton et al. 1985; Zouros & Mallet 1989; Zouros & Pogson 1994). A energia poupada por indivíduos mais heterozigotos seria então desviada para funções como crescimento e reprodução. O outro grupo de hipóteses, conhecido como sobredominância associada, assume que as aloenzimas são apenas marcadores da heterozigosidade genômica dos organismos, ou seja, a homozigosidade a nível aloenzímico estaria em desequilíbrio de ligação com a homozigosidade para genes deletérios a nível genômico. Resultados experimentais para correlação entre heterozigosidade e taxa de crescimento não são conclusivos, mostrando correlações que variam do positivo ao negativo, às vezes para a mesma espécie, como é o caso do mexilhão *Mytilus edulis* que, em quatro estudos, apresentou correlação positiva entre heterozigosidade e taxa de crescimento em dois deles, sendo que nos demais esta correlação não estava presente ou era negativa (Gaffney 1990). Estudos com o bivalve *Placopecten magellanicus*, usando marcadores moleculares de DNA nuclear (microsatélites e RFLP), não foram capazes de encontrar correlações positivas entre heterozigosidade ao nível de DNA e taxa de crescimento, o que é uma evidência contra a hipótese da sobredominância associada (Volckaert & Zouros 1989).

Apesar da correlação entre heterozigosidade e produtividade ser contenciosa, a perda de variação gênica não é desejável para as populações cultivadas. A atividade de aquicultura produz, quase que invariavelmente, uma perda de variação gênica dos estoques cultivados, o que pode ser observado em estudos com salmões (*Salmo salar* e *S. trutta*), camarões (*Penaeus japonicus*) e bivalves (*Mercenaria mercenaria* e *Crassostrea gigas*). Este fenômeno ocorre, geralmente, porque o tamanho efetivo de população (ver Quadro 1.23) envolvido em cultivos é muito menor que o aparente número de progenitores. Isto porque, o tamanho efetivo de população depende fortemente de fatores como a proporção de machos e fêmeas no estoque original e do endocruzamento ao longo das gerações. Como consequência da mudança de composição genética de estoques cultivados, o repovoamento de estoques naturais com estoques cultivados ou o escape accidental destes estoques para natureza pode ter consequências dramáticas para as populações naturais. Entre os possíveis efeitos podem ser citados: a quebra de conjuntos gênicos coadaptados e a perda de alelos exclusivos devido à introgessão ou hibridização entre os estoques natural e cultivado. Do mesmo modo, o estabelecimento de competição entre os estoques pode determinar a completa extinção de populações locais. Grandes fugas de estoques cultivados de salmões (*Salmo salar*) na Noruega, por exemplo, revelaram que a composição genética das populações atuais reflete, quase que exclusivamente, o padrão dos estoques cultivados. Alelos raros da população natural foram perdidos, bem como os seus genótipos característicos encontram-se hoje em frequências muito baixas nas populações.

A figura 1.36 resume a utilização da Genética Marinha no desenvolvimento de estratégias para aquicultura.

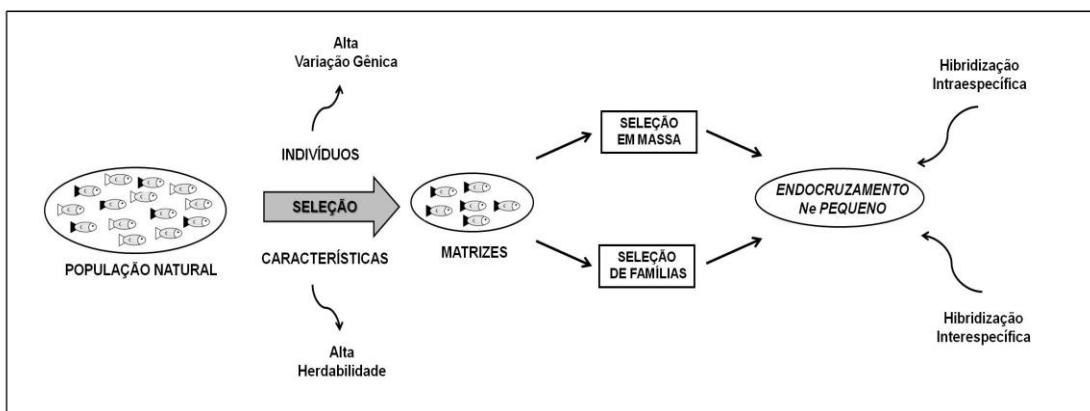


Figura 1.36. Esquema indicando a utilização da Genética Marinha Marinha no desenvolvimento de estratégias para aquicultura.

1.6.8. Bioinvasão

As invasões biológicas ou bioinvasões são a chegada, estabelecimento e subsequente difusão de espécies não nativas em comunidades naturais nas quais elas não existiam. Elas podem ser de dois tipos: expansões e introduções. As expansões consistem na dispersão natural dos organismos. As introduções, por outro lado, ocorrem quando as espécies são transportadas por atividades humanas (intencionalmente ou não) para um local onde não ocorriam. Devido ao aumento do comércio e do transporte global durante as últimas décadas, várias regiões têm sido invadidas por novas espécies (Carlton 1987; Schormann et al. 1990; Geller 1999; Holland 2000; 2001; Silva 2001a; 2004; 2012; Barnes 2002; Lee 2002; Roman & Darling 2007; Prentis et al. 2008; Duarte & Silva 2016; Pierri et al. 2016).

Na literatura sobre o assunto são encontrados termos como espécies exóticas estabelecidas e/ou naturalizadas ou invasoras. Quando uma espécie é detectada de forma recorrente em um ambiente fora da sua área de distribuição normal, apresentando indícios de aumento populacional, diz-se que esta espécie está estabelecida ou naturalizada (termos que podem ser considerados sinônimos). Uma espécie naturalizada

pode permanecer estável, com uma pequena população, durante um tempo variável até que algum fenômeno natural ou de origem antropogênica facilite o aumento da sua distribuição. Ao apresentar abundância e/ou dispersão geográfica capaz de interferir na sobrevivência de outros organismos numa determinada área ou ampla região geográfica, a espécie passa a ser considerada invasora, podendo ser enquadrada, ainda, na categoria atual ou potencial (Figura 1.37).

Um fato notável a respeito das espécies invasoras é que elas não compartilham um passado evolutivo com a comunidade invadida e isso faz das invasões biológicas um modelo interessante para compreensão dos processos que atuam na evolução dessas espécies e suas populações. As espécies invasoras estão sujeitas a, pelo menos, cinco tipos de mudanças evolutivas: *bottleneck*, hibridização, rearranjos genômicos, modificação do genoma induzida pelo estresse e efeito de determinados genes.

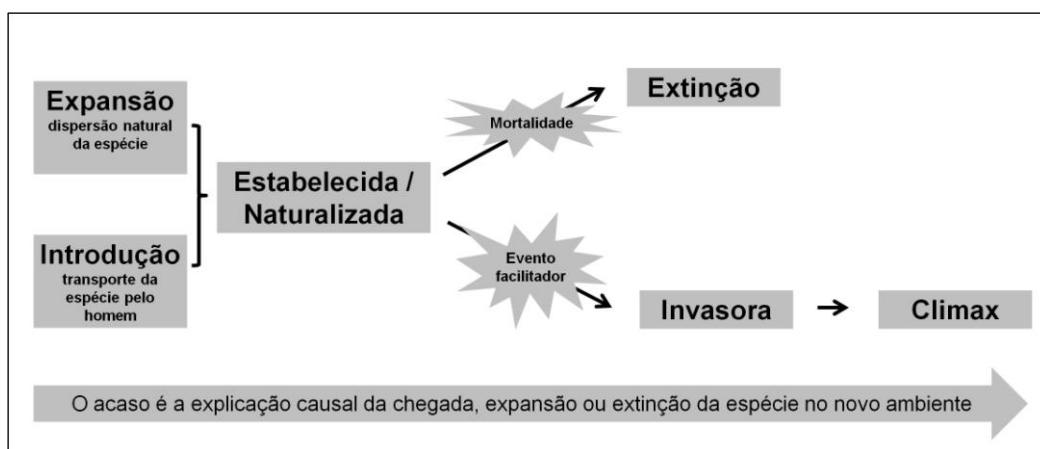


Figura 1.37. Algumas categorias definidas para as espécies exóticas, ilustrando gradiente crescente do potencial de invasão e a explicação causal da chegada, expansão ou extinção da espécie no novo ambiente.

A chegada accidental de um ou poucos indivíduos de uma espécie em um novo ambiente, é um exemplo extremo de deriva genética chamado de *bottleneck* (também conhecido como efeito fundador). Neste caso, como poucos indivíduos estão chegando ao novo ambiente, deduz-se que apenas uma pequena proporção de todos os alelos

presentes na população original estará disponível no novo ambiente, reduzindo, desta forma, a capacidade de adaptação desta população que está chegando. Da mesma forma, como esta nova população possui apenas uma amostra do conjunto de alelos da população original, pode-se imaginar que ela será, também, diferente da sua população-mãe (Figura 1.38).

Nas circunstâncias especiais da bioinvasão, outra mudança evolutiva a que estão sujeitas as espécies invasoras é a hibridização. Neste caso, ela consiste no cruzamento entre espécies de bioinvasores com espécies nativas ou com outras espécies invasoras. A hibridização, entre espécies ou entre populações da mesma espécie (dos bioinvasores com espécies nativas ou com outras espécies invasoras) pode reduzir a perda de variação gênica associada ao *bottleneck*, produzindo uma gama de novos genótipos importantes à adaptação da espécie invasora ao novo ambiente. Os efeitos positivos da hibridização ao processo de bioinvasão incluem crescimento mais rápido, maior tamanho dos híbridos e um aumento da agressividade, tornando-se uma alternativa às mudanças lentas envolvidas no processo de adaptação.

Outro tipo de fenômeno que pode determinar a adaptação rápida das populações invasoras são os vários tipos de rearranjos que ocorrem no genoma dos organismos, como poliploidia e a alloploidia. Poliplóides podem oferecer novos genótipos à ação da seleção natural e, portanto, permitir a adaptação, em curto prazo, da população invasora. É reconhecida, também, a importância de certas inversões cromossômicas na adaptação das espécies invasoras. Uma força importante na determinação dos rearranjos genômicos podem ser os transposons, que podem se inserir em diferentes posições no genoma causando um processo de liga/desliga nos genes. A exposição às condições bióticas e abióticas do novo ambiente pode, também, causar uma instabilidade no genoma, nesse caso, mediada pelo estresse ambiental. Tem sido demonstrado que

elevadas exposições aos raios ultravioletas, patógenos, bem como estresse abiótico produzem instabilidade do genoma, com o aumento da taxa de recombinação homóloga, ativação de transponse mutações. Embora alterações ao acaso do genoma mediadas pelo estresse ambiental sejam, na grande maioria das vezes, deletérias, variações gênicas benéficas associadas às pressões de seleção natural produzem adaptação. Como pode ser percebido, a Genética Marinha é uma teoria importante para entender as mudanças evolutivas envolvidas no processo de bioinvasão.

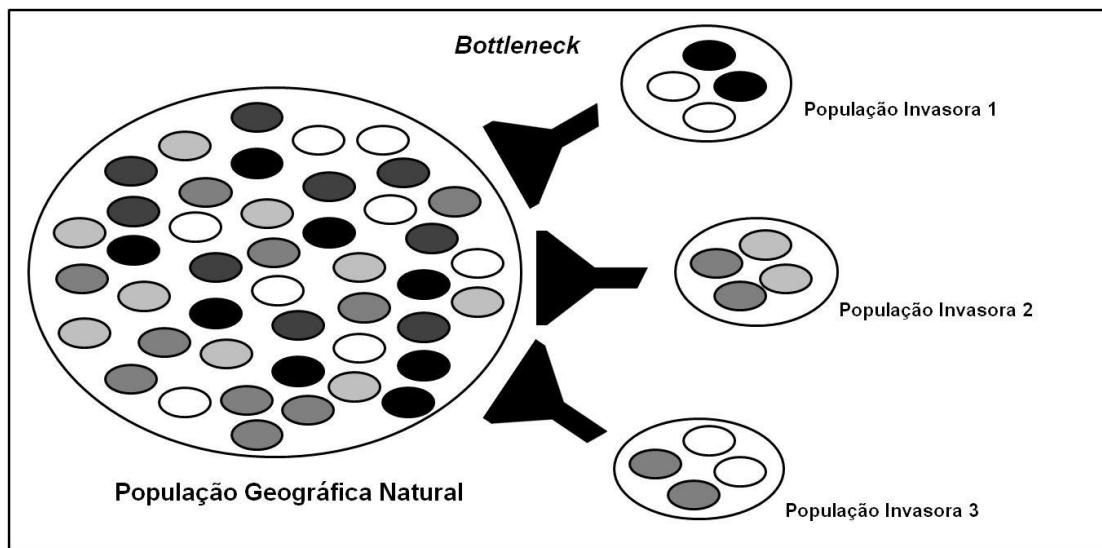


Figura 1.38. Representação gráfica de *bottlenecks* (efeito gargalo) na formação de populações invasoras.

A introdução de organismos exóticos pode ter consequências irreversíveis e devastadoras para os ecossistemas naturais, afetando tanto a biodiversidade quanto as atividades econômicas. Neste contexto, a Genética Marinha se constitui, também, numa ferramenta útil para controlar o processo global de bioinvasões. As abordagens da Genética Marinha, nesse caso, envolvem a busca e o uso de marcadores moleculares e sua interpretação usando a fundamentação teórica da genética de populações para identificação de espécies (distinguindo as espécies invasoras das nativas) e da origem das populações invasoras, bem como acompanhar a dinâmica dos processos de invasão.

No caso da identificação de espécies, a técnica de PCR tem permitido, por exemplo, a identificação de ovos, larvas e jovens no plâncton. Este tipo de abordagem é extremamente útil quando é preciso identificar as espécies invasoras nas fases mais iniciais do desenvolvimento larval, como é o caso das bioinvasões mediadas pela água de lastro. Outro exemplo é aquele das recentes investigações genéticas sobre os mexilhões da costa sul da Califórnia e da Baía de Monterey. O mexilhão *Mytilus galloprovincialis* era considerado a espécie mais abundante naquela região. Contudo, ao sequenciar uma região do gene do RNA ribossomal 16S de mexilhões do século XIX, mantidos em coleções, foi possível demonstrar que estes mexilhões eram da espécie *Mytilus trossulus*, hoje rara nesta região. Estes dados levaram a conclusão de que o declínio das populações de *M. trossulus* no sul da Califórnia e Baía de Monterey foi devido, provavelmente, à chegada do mexilhão europeu *M. galloprovincialis*, a partir de 1870. O declínio das populações nativas de *M. trossulus* não foi percebido devido à presença da espécie críptica bioinvasora (Heath et al. 1995).

No caso da origem das populações invasoras, esta pode ser inferida a partir do cálculo de índices de identidade gênica entre populações de vários possíveis pontos de origem e a população invasora em questão (ver 1.4.2.3. *Relações genéticas*). As identidades gênicas são calculadas par a par, neste caso, o par: *população invasora-possível população de origem* que apresentar o maior valor de identidade gênica estará indicando, possivelmente, a origem geográfica da bioinvasão. Identidades genéticas podem ser calculadas com base em dados de marcadores moleculares de aloenzimas, mtDNA e nDNA.

No Brasil, a Genética Marinha foi utilizada para estudos com *Isognomon bicolor*, um bivalve marinho pertencente à família Isognomonidae e endêmico do mar do Caribe (Lopes et al. 2014). Esta espécie está presente no litoral brasileiro desde

meados da década de 1980, quando foi registrado em Atol das Rocas (Natal, RN). Atualmente, este bivalve está presente em grande parte da costa brasileira, incluindo as regiões Nordeste (RN, PE e BA), Sudeste (SP e RJ) e Sul (PR e SC). Foi realizado um estudo sobre padrão variação temporal dos níveis de variação gênica desta espécie de bivalve invasor na Praia de Itaipu (Niterói, Rio de Janeiro). Os resultados revelaram consideráveis níveis de heterozigosidade em todos os anos estudados (2005, 2009 e 2013). De certa forma, estes resultados foram surpreendentes, uma vez que se esperava que as populações invasoras tivessem sofrido *bottleneck* e, portanto, apresentassem níveis de variação gênica (heterozigosidade, número de alelos etc.) baixos. Contudo, altos níveis de variação gênica em bioinvasores podem ser explicados como consequência do efeito *Wahlund* (ver 1.4.2.2. *Modelos*).

Uma marca que o efeito *Wahlund* deixa nas populações é um grande número de *loci* com desvios significativos do esperado pelo equilíbrio de Hardy-Weinberg. Numa população que é o resultado da mistura de grupos geneticamente distintos, as frequências dos alelos são diferenciadas, portanto, o equilíbrio de Hardy-Weinberg só pode ser alcançado depois de várias gerações de cruzamentos ao acaso. Antes que isto aconteça, muitas das combinações esperadas de alelos não serão encontradas na população misturada. Contudo, no estudo com *Isognomon bicolor* apenas dois *loci* apresentaram desvios significativos ao equilíbrio de Hardy-Weinberg. A estabilidade temporal dos níveis de variação gênica desta espécie na Praia de Itaipu pode ser uma indicação de que esta população tenha atingido a fase do processo evolutivo da bioinvasão conhecida como “clímax” (ampla distribuição geográfica e grandes densidades). Assim, em um primeiro momento, o efeito *Wahlund* pode ter contribuído para que a população de Itaipu respondesse às pressões ecológicas no sentido contrário à redução populacional e à extinção. Em um segundo momento, esta população

amadureceu para o “clímax” alcançando o equilíbrio de Hardy-Weinberg para maioria dos *loci*. Nesta fase, populações invasoras tornaram-se dominantes ou extremamente conspícuas, ou seja, podem ser consideradas naturalizadas.

1.6.9. Conservação

A diversidade biológica está sendo rapidamente esgotada como consequência direta ou indireta das ações humanas (poluição, extinção de *habitat*, sobrepesca, bioinvasões etc). Um número desconhecido, porém grande, de espécies já está extinto, enquanto muitas outras têm tido seus tamanhos de população reduzidos a níveis preocupantes e que as colocam em risco. Uma consequência imediata destas circunstâncias é que, hoje, muitas espécies já dependem da intervenção humana para garantir a sua sobrevivência. No ambiente marinho esta situação não é diferente e afeta, principalmente, aquelas espécies que vivem na zona costeira (e, portanto, mais sujeitas aos impactos das atividades humanas) ou que são exploradas comercialmente (espécies pescadas e ornamentais, por exemplo). Várias das aplicações da Genética Marinha discutidas até aqui têm relevância direta para os esforços de conservação da biodiversidade marinha. Por exemplo, as teorias sobre a origem e natureza da variação gênica, os estudos de estrutura de populações, a identificação de espécies crípticas, a definição de estoques pesqueiros, a identificação e acompanhamento de bioinvasões etc (Frankham 1995; Menezes & Parulekar 1998; Amos & Balmford 2001; Hedrick 2001; Solé-Cava 2001; Faleiro 2007; Solé-Cava & Cunha 2012).

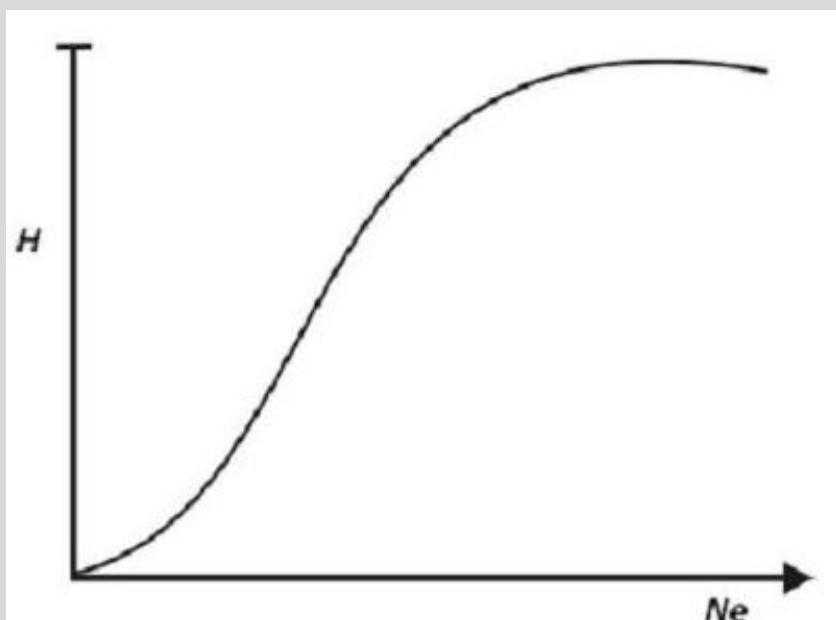
De uma maneira geral, as espécies são unidades dinâmicas e, portanto, capazes de absorver e incorporar, até certo limite, as mudanças ambientais. Ou seja, a biodiversidade abrange certo grau de resiliência que lhe permite existir no espaço e no tempo. A Genética Marinha aplicada à conservação busca compreender como os

parâmetros genéticos das populações estão relacionados com esta resiliência e, a partir disso, usá-los como informação de base para tomada de decisões. Por exemplo, dados como níveis de heterozigosidade estão diretamente relacionados com a história evolutiva das espécies (ver Quadro 1.30) e esta informação foi usada no caso dos elefantes marinhos das espécies *Mirounga angustirostris* (da costa do Pacífico da América do Norte) e *Mirounga leonina* (do sul da Argentina) (Bonnel & Selander 1974; Slade et al. 1998). A espécie *M. angustirostris* foi fortemente caçada no século XIX nos Estados Unidos, de modo que, no final daquele século, a espécie foi considerada extinta. No entanto, a partir de uma pequena colônia, de cerca de 50 casais, que habitava uma ilha desconhecida dos caçadores, as populações desse elefante marinho puderam se reproduzir e desenvolver, de modo que, no final do século XX, já chegavam a mais de 30.000 indivíduos. A despeito do grande número de indivíduos, um estudo genético da espécie revelou uma baixíssima heterozigosidade ($H < 0,2\%$) quando comparada, por exemplo, com os elefantes marinhos da Argentina ($H = 10\%$), que nunca haviam sofrido reduções populacionais devido à caça.

Uma atividade importante quando se está lidando com espécies ameaçadas é a determinação do sexo dos animais, isto por que esta informação é fundamental quando se pretende, por exemplo, formar pares para reprodução em cativeiro ou entender a estrutura familiar de uma espécie na natureza. Em animais com pouco dimorfismo sexual externo (como, por exemplo, em golfinhos) essa determinação do sexo sem produção de estresse desnecessário pode ser difícil. A Genética Marinha pode ajudar na solução desse problema a partir da análise de genes que se situam nos cromossomos sexuais. No Brasil, esse procedimento foi usado com sucesso para determinar o sexo de *Sotalia guianensis* (Cunha & Solé-Cava 2007).

Quadro 1.30. Previsões teóricas da hipótese neutralista sobre os valores de heterozigosidade em populações naturais.

Segundo os neutralistas existe uma relação direta entre tamanho efetivo de população e heterozigosidade. Isto porque em populações grandes a variação que está entrando por taxa mutação demora a se perder por efeito de deriva, que é pequena em populações grandes. Por outro lado, como em populações pequenas o efeito de deriva é elevado, as mutações que entram se perdem rapidamente (ver Figura 1.4 seção 1.2.1.3. *Controvérsias evolutivas*). Assim, heterozigosidades altas são esperadas em populações grandes e baixas heterozigosidade em populações pequenas. Esta relação entre heterozigosidade e tamanho efetivo de população pode ser estimada facilmente por meio de uma fórmula simples (ver Quadro 1.23 em 1.6.1. *Variação Gênica*). Desta forma, para uma dada taxa de mutação, é possível construir uma curva teórica que representa a distribuição esperada da variação gênica (medida como heterozigosidade) em função dos Ne's, como pode ser observado abaixo no gráfico.



Esta é a teoria que serve como base para o uso dos dados de Genética Marinha nas intervenções de manejo de populações de tamanho pequeno usando dados de heterozigosidade.

A Genética Marinha pode ter um papel importante, também, na conservação dos ambientes marinhos. Por exemplo, foi demonstrado, recentemente, a partir do sequenciamento de DNA mitocondrial e nuclear, que as espécies de corais dos recifes do Brasil e do Caribe formam, ao contrário do que se pensava, um grupo evolutivo original e bem diferenciado dos recifes do Oceano Pacífico. Desta forma, no caso de

políticas de conservação, a região dos recifes de corais do Atlântico, incluindo o Brasil, passa a ser tão importante quanto a do Oceano Pacífico.

A criação, manutenção e manejo de áreas marinhas protegidas é outro exemplo da utilização da Genética Marinha para o problema da conservação. O pressuposto para criação e manutenção dessas áreas é aquele de que este tipo de unidade de conservação é capaz de aumentar, estabilizar e preservar os ecossistemas marinhos, não só na região das reservas como, também, para além dos seus limites. Portanto, um dado importante para entender o funcionamento e a dinâmica dessas áreas é o quanto populações espalhadas ao longo da faixa de distribuição de uma espécie estão ligadas por meio do movimento de larvas, recrutas, juvenis ou adultos, ou seja, o seu nível de conectividade (Palumbi 2001). Nestes casos, contudo, é importante se ter em mente que a utilização dos dados da Genética Marinha deve ser feita com cuidado, uma vez que eles são mais bem interpretados em uma escala evolutiva do que em uma escala ecológica (ver **1.4.2. Variação Interpopulacional**).

Para além da biologia pura e simples, os dados de conectividade do ambiente marinho podem ser importantes, também, para problemáticas sociais relacionadas com alguns tipos de áreas marinhas protegidas. Este é o caso, por exemplo, das unidades de conservação conhecidas como Reservas Extrativistas (Resex). A principal diferença das Resex em relação a outras como, por exemplo, os “Santuários Ecológicos” é que, nas primeiras, as populações humanas são consideradas como um componente do ecossistema e, portanto, as suas regras de gestão são feitas “pela” própria comunidade e não “para” a comunidade. Antropólogos têm relatado que, nas Resex, a definição dos recursos naturais como ‘locais’ é um importante componente na problemática da identidade social de uma Reserva extrativista marinha. Ou seja, os recursos “locais” são

aqueles para os quais se tem uma representação tradicional sobre sua ocorrência e abundância e, portanto, re-affirmam a identidade local de certas comunidades.

Em Arraial do Cabo, no estado do Rio de Janeiro, foi criada, em 1997, uma Resex. A discussão identitária nesta reserva incluiu vários atores sociais, entre eles os marisqueiros, que tentavam incorporar seus interesses na gestão da reserva. Contudo, muita discussão se dava quanto a natureza do recurso natural por eles explorado, o mexilhão *Perna perna*. Tanto assim que, hoje, em Arraial do Cabo, os marisqueiros estão organizados na ACRIMAC (Associação dos Coletores e Criadores de Mariscos de Arraial do Cabo), diferente da AREMAC (Associação da Reserva Extrativista Marinha de Arraial do Cabo). Um estudo sobre a natureza do recrutamento deste mexilhão nos costões de Arraial do Cabo demonstrou que as larvas da espécie ao invés de serem apparentadas com os adultos do local (o que caracterizaria um recrutamento local), eram distantes genéticamente deles e entre si também (caracterizando um recrutamento caótico). Um resultado interessante no desenvolvimento das discussões identitárias na Resex Mar de Arraial do Cabo.

Ainda com relação ao mexilhão *Perna perna*, ele é fonte relevante de renda para famílias que, no Brasil, realizam tanto a seu extrativismo quanto o seu cultivo. Contudo, atualmente vem sendo discutida a hipótese de que ele seria uma espécie exótica do litoral do Brasil, tendo sido introduzida no país há pouco mais de 400 anos, devido ao tráfego de navios negreiros oriundos do continente africano. Esta hipótese é baseada, fundamentalmente, em dados de zooarqueologia que não revelam a presença desta espécie em sambaquis tão antigos quanto 1.000-8.000 anos A.P. Esta discussão foi responsável pela proibição do cultivo da espécie em alguns estados do país. Nesse contexto, a Genética Marinha tem se envolvido em um esforço colaborativo com outras disciplinas como a arqueologia, biologia molecular, física e biogeografia, entre outras

(Souza et al. 2003; 2004; 2009; Silva & Souza 2004; Weber & Silva 2008; Fernandes et al. 2011). Contudo, neste caso, a contenda não está ainda resolvida. Ah! Fala-se de 2016.

Outra aplicação da Genética Marinha para conservação das espécies está relacionada ao controle da pesca ilegal. Algumas espécies que estão ameaçadas de extinção são protegidas por lei, no entanto, ainda assim, são pescadas de forma irregular. Para esconder o delito, os contraventores processam a carne das espécies protegidas para dificultar a sua identificação pelas autoridades fiscalizadoras. Esta é, por exemplo, a situação da caça realizada sobre a maior parte das espécies de baleias e que é proibida por lei. Contudo, como algumas espécies podem, ainda, ser caçadas, como a baleia Minke, que é considerada abundante, os baleeiros (embarcações especializadas na caça de baleias) que caçam a Minke matam, também, outras espécies que tem a caça proibida. Conhecimento das sequências de diferentes espécies da baleias permitiram que, no Japão, filés apreendidos pudessem ter as suas espécies de origem identificadas e, os infratores, punidos.

A mesma estratégia foi usada no Canadá com relação à pesca de vieiras. As vieiras *Placopecten magellanicus* (espécie maior) e *Chlamys islandica* (espécie menor) são moluscos bivalves comercialmente importantes e têm a sua pesca regulada pelo tamanho dos indivíduos (Kenchington et al. 1993). Contudo, após o processamento realizado dentro das embarcações, o controle fica difícil. Músculos adutores destas duas espécies foram analisadas através de três técnicas (aloenzimas, sequenciamento e RFLP) com o objetivo de definir marcadores genéticos específicos para as duas espécies. Os resultados evidenciaram que todas as três técnicas podem ser utilizadas de modo efetivo nos programas de fiscalização da pesca de vieiras. Os resultados de eletroforese de aloenzimas evidenciaram dois *loci* diagnósticos (ver **1.6.3. Sistemática e**

Taxonomia). Os resultados de sequenciamento do gene 18S rRNA mostrou 15 diferenças de nucleotídeos (em mais ou menos 1815 pares de base). Com relação ao RFLP, a digestão com a enzima de restrição Xho I evidenciou que a espécie *Placopecten* apresentava um padrão de duas bandas e, a espécie *Chlamys*, três (ver 4.3.6.3. RFLP).

1.6.10. Biogeografia

Os estudos em biogeografia marinha são, essencialmente, descritivos e comparativos e, na sua fase atual, encontram-se, muito ainda, preocupados em caracterizar padrões de distribuição geográfica e delimitar unidades biogeográficas. As unidades biogeográficas são definidas pelo nível de endemismo e descontinuidades de *habitat* reconhecidos. Por exemplo, uma província biogeográfica é, geralmente, definida por um nível de endemismo de 10%. Outros tipos de unidades biogeográficas são as regiões e os centros de endemismo (Briggs 2006; Floeter et al. 2009; Pelc et al. 2009; Miranda & Marques 2011; Miller et al. 2013).

A preocupação da biogeografia marinha em definir unidades biogeográficas advém, em grande parte, do fato de que o ambiente marinho é percebido como homogêneo e os *taxa* marinhos como bons dispersores. Como já foi visto neste capítulo, as coisas não são bem assim. Em muitas regiões do mundo foram identificados limites importantes a distribuição das comunidades marinhas como as correntes oceânicas, gradientes térmicos e descontinuidades de *habitat*. Mais que isso, a extensão destes limites é influenciada, também, por fatores como traços da história de vida e ecologia das espécies.

De um modo geral, os processos envolvidos na determinação dos padrões de distribuição espacial dos *taxa* marinhos podem ser recentes ou históricos. Os processos

recentes são, geralmente, de ordem ecológica como, por exemplo, temperatura, salinidade, disponibilidade de alimento e sedimentos e afetam, principalmente, o nível infra-específico, ou seja, como se estruturam as populações dentro de uma espécie. Exemplos de como a Genética Marinha pode auxiliar a entender estes padrões foram descritos em **1.6.2. Estruturação**. A escala de tempo, nestes casos, é curta e a biogeografia preocupada com estes padrões e processos é chamada de biogeografia ecológica.

Nos casos nos quais os padrões de distribuição dos *taxa* são determinados por processos de grande escala temporal, a biogeografia é chamada histórica. Os níveis estudados são, então, supraespecíficos (espécies diferentes, gêneros, família etc.) e os principais processos são de ordem evolutiva (especiação, extinção, irradiação) e geológica (tectônica de placas). Tanto no caso da biogeografia ecológica quanto da biogeografia histórica, os dois grandes modelos explicativos dos padrões biogeográficos são a dispersão e a vicariância.

Se, como já foi dito, a biogeografia marinha é, fundamentalmente, uma ciência descritiva e comparativa, o auxílio da Genética Marinha tem possibilitado a ela que os seus modelos explicativos sejam testados experimentalmente. Por exemplo, no caso da dispersão, a espécie *Donax deltoides* foi usada para testar a influência de correntes oceânicas no sudeste da Austrália. Esta espécie é um molusco que possui um grande potencial de dispersão por meio das suas larvas planctotróficas. Além disso, não possui restrições de *habitat* na região, sendo, por isso, um bom modelo para entender a influência das correntes na sua dispersão naquela região. Foram analisadas as frequências alélicas de dez *loci* polimórficos de microsatélites e sequências de mtDNA em 11 localidades de uma região de convergência de correntes. Os dados de mtDNA não forneceram nenhuma evidência de estruturação genética, de acordo com a

expectativa de grande dispersão da espécie mediada pelas correntes. Do mesmo modo, os resultados de microsatélite indicaram uma estrutura genética consistente com as principais correntes oceânicas da região, evidenciando que elas são o fator mais influente no padrão de distribuição das populações *D. deltoides* no sudeste da Austrália.

A biogeografia marinha tem se beneficiado também dos estudos de filogeografia. Quebras filogeográficas podem refletir impedimentos históricos ou atuais ao fluxo gênico. A congruência destas quebras em várias espécies fornece uma visão histórica evolutiva da conectividade genética entre as populações. Contudo, examinar a concordância de quebras filogeográficos não é simples, pois as escalas de amostragem variam em estudos de genética de populações e os organismos utilizados possuem diferentes histórias de vida. A acumulação de dados de marcadores moleculares (aloenzimas, microsatélites, mtDNA) tem começado a revelar alguns padrões interessantes como, por exemplo, de que a estrutura filogeográfica em espécies com larvas planctônicas parece refletir padrões oceanográficos contemporâneos, enquanto em espécies com dispersão restrita, a estrutura parece refletir processos históricos.

Um estudo com Pleurobranchomorpha que são lesmas do mar e representam um clado de gastrópodos marinhos distribuídos por todo o mundo teve a sua biogeografia investigada em escala global (Göbbeler & Klussmann-Kolb 2010). As espécies deste clado apresentam uma morfologia bastante homogênea e com vários caracteres diagnósticos. Contudo, sua filogenia não é bem estabelecida, o que deixa mal resolvidos muitos aspectos da sua biogeografia. Por exemplo, na última classificação realizada para o clado, ficou definido que ele é constituído de uma única superfamília (Pleurobranchoidea) com uma única família (Pleurobranchidae). Esta família é dividida em duas subfamílias. A primeira subfamília Pleurobranchaeinae compreende três gêneros: *Pleurobranchaea*, *Euselenops* e *Pleurobranchella*. A segunda subfamília

Pleurobranchinae é dividida em três tribos: Pleurobranchini, Berthellini e Bathyberthellini. A tribo Pleurobranchini é constituída pelo gênero *Pleurobranchus* (o mais diverso de Pleurobranchomorpha). A tribo Berthellini compreende os gêneros *Berthella*, *Berhellina*, *Pleurehdera* e *Tomthompsonia*. Por fim, a tribo Bathyberthellini inclui os gêneros *Bathyberthella* e *Boreoberthella* (ver a figura 1.39 que apresenta uma estruturação taxonômica dos *taxa* que foram referidos neste parágrafo).

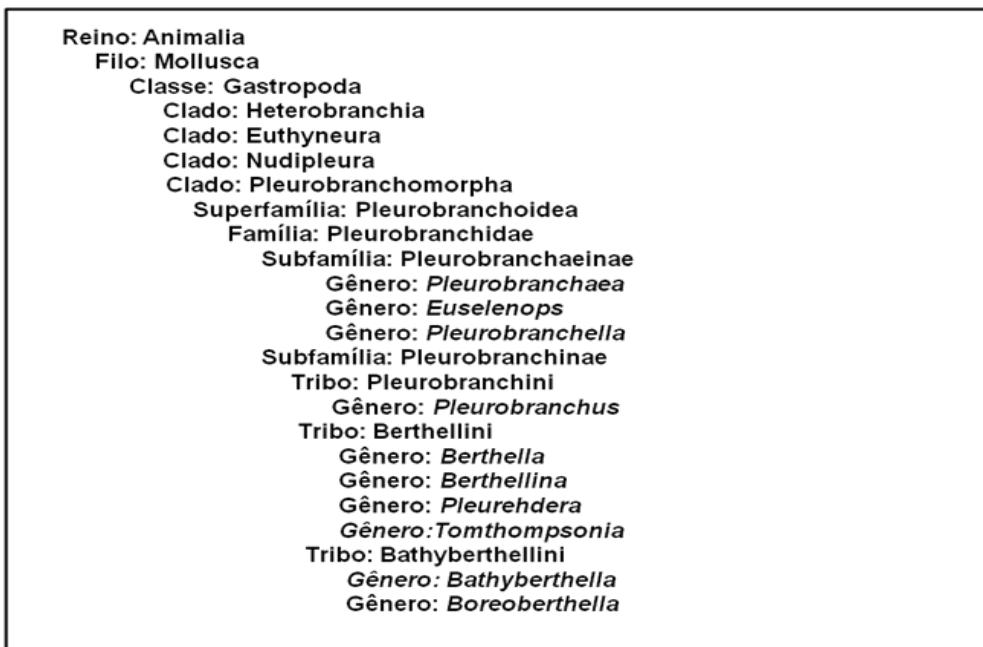


Figura 1.39. Estruturação taxonômica dos *taxa* de Pleurobranchomorpha. Os sete gêneros que foram amostrados no trabalho descrito são: *Pleurobranchaea*, *Euselenops*, *Pleurobranchus*, *Berthella*, *Berhellina*, *Tomthompsonia* e *Bathyberthella*.

Para entender a origem geográfica do clado Pleurobranchomorpha foram realizados estudos com quatro marcadores moleculares (os *loci* 18S rDNA, 28S rDNA, 16S rDNA e COI) que foram amplificados em reações de PCR e, depois, sequenciados. Foram 36 *taxa* estudados que correspondiam a 17 espécies de Pleurobranchomorpha (abrangendo sete dos dez gêneros descritos), 16 espécies representantes de todos os principais subgrupos da Ordem Opisthobranchia, duas espécies do clado Heterobranchia e *Littorina littorea* como grupo de fora.

As análises filogenéticas apontaram para monofilia de Pleurobranchomorpha, entretanto, não para os principais subgrupos Pleurobranchaeinae e Pleurobranchinae. Mais interessantes que tudo, as análises indicaram uma origem antártica para Pleurobranchinae. As estimativas de tempos de divergência, baseadas no relógio molecular, indicaram que o clado Pleurobranchomorpha é muito antigo, enquanto Pleurobranchinae divergiu no inicio do Oligoceno e depois sofreu uma rápida radiação durante Oligoceno e início do Mioceno. Nesta irradiação, *Tomthompsonia* divergiu e saiu da Antartica. Esta interpretação dos dados da filogenia molecular coincide com dois grandes eventos geológicos: i) início da glaciação na Antártica e ii) abertura da passagem de Drake que permitiu a formação da Corrente Circumpolar Antártica (ACC). Estas mudanças súbitas e dramáticas de clima levaram ou à migração posterior do último ancestral comum dos Pleurobranchinae restantes para regiões mais quentes, enquanto a ACC pode ter contribuído para a dispersão das larvas da espécie para o Atlântico Oriental. Na figura 1.40 este cenário biogeográfico esta representado. Como pode ser percebido a Genética Marinha oferece um “auxílio luxuoso” aos estudos da biogeografia marinha.

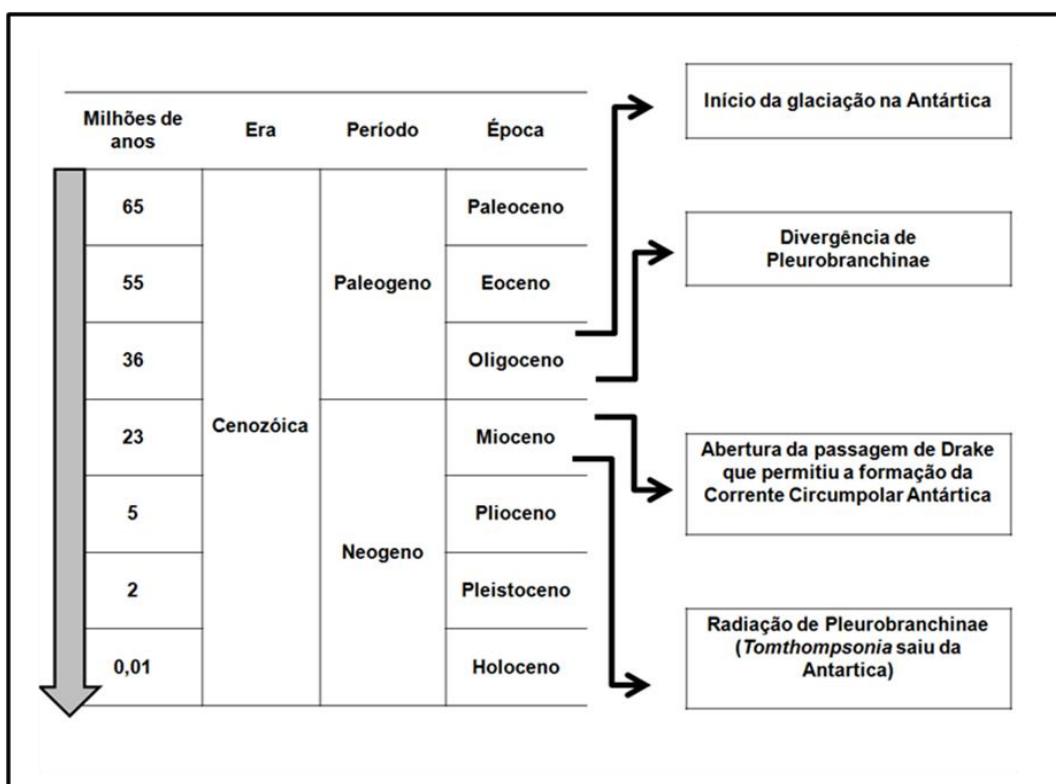


Figura 1.40. Tempos geológicos com indicação dos períodos de divergência e radiação de Pleurobranchinae.

1.6.11. Microbiologia

A microbiologia é o ramo da biologia que estuda os microrganismos em sentido amplo, incluindo vírus, bactérias, archaea e eucariontes unicelulares. Nos oceanos, para além da biodiversidade discutida até aqui, existe uma vasta vida microscópica. Estudos de filogenia estimaram que a diversidade microbiana no mar pode variar desde centenas até milhares de espécies diferentes por mililitro de água. Mais do que a sua diversidade, a vida microbiana é essencial em praticamente todos os processos biogeoquímicos e, dessa forma, tem um impacto direto no clima e na vida macroscópica da Terra (DeLong & Karl 2005; Heidelberg et al. 2010).

O oceano é um sistema integrado, onde transformações de matéria e energia são regidas por processos físicos e químicos interdependentes, mas, também, e, talvez, principalmente, por processos bióticos regidos pelos microorganismos. Contudo, se os

fundamentos da física e da química do oceano estão bem estabelecidos, no caso da microbiota, a própria diversidade é desconhecida. Portanto, a despeito da sua relevância, a compreensão sobre o metabolismo, ecologia e evolução da vida microbiana marinha é precária. A Genética Marinha tem estimulado avanços nesta área, principalmente pela incorporação das plataformas NGS que têm revolucionado a forma de se obter informações genéticas sobre as comunidades microbianas.

De uma maneira geral, a forma de catalogação de espécies em microbiologia se baseia na tomada de amostras ambientais e no cultivo das espécies em diferentes meios de cultura (liquida e sólida). Contudo, menos de 1% das espécies marinhas é suscetível a estes cultivos, o que sempre tornou a sua identificação difícil. Pesquisas de sistemática molecular têm circunscrito este problema e estão revelando uma imensa diversidade, acima de todas as expectativas. Desta forma, mesmo depois de algum tempo de aplicação das técnicas moleculares ao estudo da biodiversidade microbiana, a descoberta de novas espécies continua sendo comum. Investigar a diversidade microbiana por meio da Genética Marinha é feito a partir da extração de DNA e RNA diretamente de solos e águas marinhas, a submissão destes DNA's e RNA's a reação em cadeia da polimerase, o seqüenciamento direto dos produtos de PCR e a comparação das sequências obtidas com bancos de sequências. Sequências ausentes no banco de dados representam novas espécies.

Além da análise de sistemas microbianos livres, ferramentas moleculares têm desempenhado um papel importante na elucidação de relações endosimbióticas. Análises baseadas em rRNA foram usadas para identificar bactérias associadas com hospedeiros invertebrados que vivem em fontes hidrotermais. O foco destas abordagens tem sido tentar compreender a especificidade das relações ecológicas e evolutivas entre simbionte-hospedeiro. Um trabalho desta natureza foi realizado, por exemplo, em

bactérias luminescentes simbióticas que habitam os órgãos de luz de peixes e lulas. O objetivo era investigar a ecologia do simbionte enquanto não associado com o seu hospedeiro. Os resultados demonstraram que a abundância e distribuição de uma bactéria marinha (*Vibrio jscheri*) na coluna de água era diretamente influenciada pela presença de seu hospedeiro animal (a lula, *Euprymna scolopes*) no ambiente.

Os micróbios eucariontes compreendem um conjunto diversificado de criaturas fototróficas e heterotróficas que são conhecidos por desempenhar papéis fundamentais em processos ecológicos. Os protistas, por exemplo, são responsáveis pela maior parte da conversão de produtividade bacteriana. No entanto, a contribuição desses organismos para o funcionamento dos sistemas marinhos é, muitas vezes, negligenciado. Além disso, neste caso, o uso da abordagem genética é limitado, uma vez que os organismos eucariontes têm genomas maiores (10-10.000 vezes maior que o genoma de procariontes) sendo, portanto, mais difícil de obter, para eles, seqüenciamentos de genomas inteiros.

Superando a dificuldade do trabalho com eucariontes, um estudo sobre biodiversidade e sucessão sazonal do grupo das haptophyta (também conhecidas como "algas douradas/amarronzadas") foi realizado a partir do sequenciamento parcial da região 18S do rDNA. Amostras mensais foram coletadas durante dois anos em uma estreita baía na área de Skagerrak (entre Noruega, Suécia e Dinamarca). A região possui um ciclo anual de temperatura que varia de 1°C até 18°C. Os resultados evidenciaram uma sucessão sazonal marcante com cerca de 30% da variação sendo explicada pelo ciclo anual de luz/temperatura, com a riqueza sendo mais alta no outono e mais baixa na primavera. Muitas das espécies encontradas neste estudo não tinham correspondência com espécies conhecidas, evidenciando mais uma vez a biodiversidade escondida nas assembléias microbianas e que tem sido revelada pela abordagem da Genética Marinha.

Estes foram apenas poucos exemplos da utilização da abordagem da Genética Marinha aplicada a microbiologia marinha. Certamente que o maior impacto desta abordagem tem sido na identificação de novas espécies. Contudo, o futuro reserva, ainda, novas e importantes aplicações desta disciplina na compreensão do papel e função dos microorganismos marinhos.

1.7. CONCLUSÃO

Neste capítulo buscou-se apresentar, descrever e ilustrar com exemplos, as mais diversas áreas nas quais a Genética Marinha pode ser aplicada. Não foi a intenção apresentar uma revisão detalhada da bibliografia ou, ainda, colocar os problemas na perspectiva das últimas inovações técnicas. Pelo contrário, foi feita uma opção deliberada de apresentar as bases históricas, empíricas, conceituais e científicas da disciplina, fazendo para àqueles que trabalham ou se interessam por genética e evolução, biologia molecular e biologia marinha uma apresentação. Leve e ilustrativa o suficiente para um leigo curioso ou um estudante interessado; precisa e informada o suficiente para um profissional ou colega da área. Numa perspectiva ao mesmo tempo geral e específica. Não se teve a esperança dos complementares, mas aceitou-se o risco dos contrários e o movimento da sua tensão, numa dialética materialista.

No geral, esperava-se que o leitor encontrasse a fundamentação teórica e empírica da genética de populações e dos marcadores moleculares. No específico, o texto tentou discutir como os modelos se aplicam ao ambiente marinho e sua biodiversidade. Nos oceanos, os padrões de biodiversidade têm sido descritos em duas categorias. Primeiro, a produção alopátrica de espécies endêmicas em regiões nas quais barreiras geográficas (tais quais os padrões globais de temperatura) propiciam o isolamento e, segundo, a especiação simpática em centros de origem, com sua posterior dispersão para outras áreas. Contudo, grandes padrões não explicam nem respondem detalhes. Neste sentido, o texto tentou enfrentar, também, as ideias em torno do macro e do micro. Desta forma, com geral e específico, macro e micro, a decisão de escrever este capítulo foi perigosa, por isso buscou-se, sempre que possível, a história das idéias e sua dialética de teses, antíteses e sínteses, mas sem verdades.

Como foi dito no começo, a Genética Marinha tem um espaço de existência que é uma interseção, ou seja, a zona sombreada do cruzamento entre genética de

populações (teoria), biologia molecular (técnicas) e biologia marinha (objetivos). Portanto, neste entremeio, o “topos” desta disciplina se configura, de fato, como um “utopos”. Sem querer dizer que a Genética Marinha não tem um lugar ou, que como uma utopia, ela não existe. O que esta se dizendo é que o *status epistemológico* desta disciplina é algo a se pensar. Esta questão foi uma provocação deste capítulo, desde a sua introdução (na qual o problema foi explicitado), passando pela sua própria estrutura, até onde temos que acabar. Contudo, não se quis resolvê-la teoricamente, embora se espere que a heurística da Genética Marinha tenha sido demonstrada de forma clara e sonante.

Neste ponto, portanto, o leitor poderia se perguntar, talvez, que, então, este capítulo pudesse ser re-nomeado como “Genética”, simplesmente. Ou seja, um texto ilustrando técnicas, teorias e casos de transmissão das características biológicas de geração a geração. É preciso que se diga que não seria um erro absoluto, afinal muitos dos estudos fundadores da Genética Marinha revelavam a engenhosidade metodológica dos precursores e seu esforço de demonstração teórica da herança biológica. Contudo, o leitor, possivelmente, concordaria que ficaria faltando água neste moinho.

A água que estaria faltante é a perspectiva populacional da herança biológica e, portanto, sua inevitável ideia de mudança. Ou seja, esta genética é água em movimento, flui em mudança no tempo e no espaço, é evolução! O leitor poderia, então, concluir: este capítulo poderia, portanto, se chamar “Genética e Evolução” (“Genética e Água”, talvez) ou, simplesmente, “Evolução”, já que a água da genética que se está falando é a evolução. Mas, ainda assim, estaria faltando alguma coisa: toda água em movimento, flui para o mar!

O que faz da água mar e do mar oceano é sal, mais que cor, mais que extensão. Assim, a genética e a evolução de que se tratou aqui têm água, têm movimento e têm

sabor! Um sabor marisma, estuário, laguna, costeiro, mas, também, ilha oceânica, mar profundo, antártica, ao largo. É de fato uma Genética Marinha, se não no lugar (epistemológico), na profissão (heurística). É neste sentido que este capítulo tratou de uma Genética Marinha. Ou seja, ele se dedicou a água e ao sal da genética.