Title: EXTENT OF DNA HOMOLOGY IN SOME DISTANTLY AND
INTIMATELY RELATED FISHES

An attempt was made to differentiate between runs of steelhead by comparing the reassociations of their DNA. Conditions for the reassociation were studied using *E. coli* DNA, and observations were made spectrophotometrically. Of the five buffers examined, 1.0 molar sodium perchlorate plus one-hundredth molar tris-HCl, pH 7.1, at 60° or 70°C produced the fastest reassociation rates of 40,000 psi sheared DNA, over ten times faster than that of 0.12 molar sodium phosphate, pH 6.8. The rate increase was necessary for observation of the very slowly reassociating DNA.

The genomes of the chinook jack, steelhead, and starry flounder were characterized and compared with respect to the composition of repeated sequences, similarity in base sequencing, and the amount of information in the nonrepeated sequences. The steelhead has a large, highly repeated fraction, sequences repeated at many intermediate frequencies, and a nonrepeated portion comprising...
20-40% of the total and having a complexity greater than $2 \times 10^9$ nucleotide pairs. Chinook DNA was quite similar but had distinct differences in the base sequencing of the less repeated region. Flatfish DNA differed greatly from that of the steelhead. The amount of the nonrepeated DNA was two thirds of the total while the complexity was only $1 \times 10^9$. There was no significant amount of DNA repeated at intermediate frequencies, and there were no measurable sequence similarities.

The DNA from the Rogue River, the Siletz River, and Sashin Creek steelhead had no differences measurable by the techniques used.
Extent of DNA Homology in Some Distantly and Intimately Related Fishes

by

Anthony John Gharrett

A THESIS

submitted to

Oregon State University

in partial fulfillment of the requirements for the degree of

Master of Science

June 1973
APPROVED:

Redacted for Privacy
Professor of Fisheries

Redacted for Privacy
Head of Department of Fisheries and Wildlife

Redacted for Privacy
Dean of Graduate School

Date thesis is presented August 16, 1972.
Typed by Velda D. Mullins for Anthony John Gharrett.
ACKNOWLEDGEMENTS

I am indebted to Dr. R. C. Simon, Leader of the Oregon Cooperative Fishery Unit, for his support, patience and guidance throughout this study. I also appreciate his editorial comments on this thesis. For reviewing my thesis, I wish to thank Dr. G. S. Beaudreau and Dr. C. E. Bond, both of Oregon State University.

I appreciate the helpful advice concerning DNA reassociation methodology and interpretation given by Dr. R. J. Seidler of Oregon State University and Dr. R. J. Britten of the Carnegie Institute of Washington, Department of Terrestrial Magnetism.

This investigation was supported by the Oregon State Game Commission, Project Number AFS-28.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II METHODS AND MATERIALS</td>
<td>5</td>
</tr>
<tr>
<td>DNA Preparation</td>
<td>5</td>
</tr>
<tr>
<td>Melt Curves</td>
<td>6</td>
</tr>
<tr>
<td>Sheared DNA Size</td>
<td>7</td>
</tr>
<tr>
<td>Reassociation of DNA</td>
<td>8</td>
</tr>
<tr>
<td>Calculation of Reassociation Kinetics</td>
<td>11</td>
</tr>
<tr>
<td>III RESULTS AND DISCUSSION</td>
<td>14</td>
</tr>
<tr>
<td>Buffer Comparisons</td>
<td>14</td>
</tr>
<tr>
<td>The Steelhead Genome</td>
<td>19</td>
</tr>
<tr>
<td>Comparison of Steelhead and Starry Flounder DNA</td>
<td>22</td>
</tr>
<tr>
<td>Chinook-Steelhead Comparisons</td>
<td>25</td>
</tr>
<tr>
<td>Perspectives</td>
<td>31</td>
</tr>
<tr>
<td>LITERATURE CITED</td>
<td>33</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>36</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Estimates of maximum relative decrease in absorbance of sheared E. coli DNA from the best least squares fit to the straight line $A(0) - A(t)$ on time (t) obtained by varying $A(\infty)$.</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>Properties of the straight line portion of the plot of percent reassociation on log C t of E. coli DNA reassociated in different buffers and at different temperatures.</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Second order rate constants $\left(\frac{\text{liter}}{\text{mole} \cdot \text{second}}\right)$ for the reassociation of E. coli DNA.</td>
<td>21</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reassociation of <em>E. coli</em> DNA sheared and reassociated in the buffers SSC, 2SSC, PB, HClO4, and ClO4</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>The reassociation of steelhead DNA</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>The rate of reassociation of steelhead DNA</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>The reassociation of the DNA's of starry flounder, Rogue steelhead, a mixture of equal parts of both, and a theoretical curve for the mixture of flounder and steelhead DNA</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>The slope of the reassociation curves shown in Figure 4 plotted against log C° t.</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>DNA reassociation curves of chinook jack, Siletz steelhead, and an equal parts mixture of the two.</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>DNA reassociation curves of HAP fractionated fast reassociating DNA from chinook jack, steelhead, and an equal mixture of the two.</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>DNA reassociation curves of HAP fractionated slowly reassociating DNA from chinook jack, steelhead, and a mixture of equal amounts of the two.</td>
<td>26</td>
</tr>
<tr>
<td>9</td>
<td>DNA reassociation from Sashin Creek steelhead, Rogue River steelhead, and a mixture of equal amounts of the two.</td>
<td>26</td>
</tr>
<tr>
<td>10</td>
<td>Theoretical second order curves for different rate constants (k), and two theoretical two-component curves.</td>
<td>42</td>
</tr>
</tbody>
</table>
EXTENT OF DNA HOMOLOGY IN SOME DISTANTLY AND INTIMATELY RELATED FISHES

INTRODUCTION

Management policies applied to one population of anadromous fish are not necessarily realistic when applied to another population. Management is implemented by changing the environment in specific ways. Because of variation in the genetic compositions and environments of different populations, there is no reason to expect that a measure applied to one population will produce the same results on another population. For the most effective management, therefore, it is necessary that one be able to distinguish between and, hence, make decisions on the basis of breeding populations. The steelhead (Salmo gairdneri) is one species that would benefit from more extensive knowledge of the size of and overlap between--both temporal and spatial--breeding populations.

The fragmentation of organisms into discrete populations is a result of geographical and behavioral isolation. The genetic information of isolated populations is subject both to different selection pressures and to random drift. Therefore, it is probable that differences exist in the genetic composition of groups where interchange of adults is restricted. Genetic composition of a "group" or population is determined by the base sequencing of the DNA of its
individuals. Variations in these sequences result from base changes, deletions, and duplications. Detection of these differences is the most basic way to determine the existence of discrete populations and the extent of their dissimilarity. This comparison also eliminates the obscuring influence the environment has in most other comparisons. However, fish populations have never been examined previously with the intention of attempting to measure population-level differences in nucleic acids.

Double-stranded DNA may be dissociated into single strands and, under appropriate conditions, may be reassociated to reform a duplex displaying high specificity in base pairing [16, 19]. Some double strands in this duplex may be defined as a "hybrid" where DNA of different species or populations are mixed. Hybrid formation, when sufficient complementarity between heterologous strands exists, enables an estimation of base-sequence similarity of the DNA from the two sources [12].

Various techniques have been used for the comparisons of DNA sequences. Two basic categories of comparison procedures are those in which DNA is reassociated in solution [20] and those in which single stranded DNA of one organism, fixed in agar [12] or on nitrocellulose filters [9, 25] to prevent reassociation, is incubated with short, dissociated strands of a second organism. In the latter techniques (as opposed to solution studies) only the
hybrids are measured, the interference of reassociation of homologous DNA sequences is eliminated. While reassociation in solution follows kinetics that approximate a second order reaction [22, 26], the kinetics of fixed DNA are uncertain. In addition, the times required for extensive reassociation of DNA of highly complex animals are prohibitively long [5]. Two excellent review articles considering both the theory and methodology of reassociation and hybridization of DNA have been written by P. M. B. Walker [24] and by Kohne and Britten [11].

The progress of reassociation in solution may be measured most practically by observing the decrease in absorbance at 260nm accompanying duplex formation or by the separation of double and single strands on hydroxyapatite [1]. While the latter method measures strand pairing and is an excellent method for fractionating the DNA with respect to the repetitiveness of the sequences, the former technique measures the base pairing of the reassociation and is, therefore, more likely to discriminate small numbers of base changes. Spectrophotometry provides a more frequently used accurate, and convenient means of measurement.

All organisms are typified by a genome which is defined as the complement of nucleotide-base sequences contained in a haploid chromosome set. In higher animals (fishes included) some sequences are repeated a great many times while others appear to
exist as unique, single-copy sequences. In addition, within the same genome, there exists a gradation in the extent to which a sequence is duplicated, approaching about one million copies in the extreme \([4, 5]\). The total amount of nonrepeated DNA in the genome is the maximum amount of information the bases code. This amount of DNA is the "complexity" of the organism.

The second order nature of DNA reassociation indicates that the repeated sequences will reassociate earlier than the nonrepeated sequences because of the greater probability of collision of the more abundant sequences. Analysis of the reassociation curves will, therefore, provide information on the relative amounts of repetition present within the genome \([5]\).

In this study, the degree to which comparisons of DNA may be used to differentiate some selected fishes was examined. Since DNA from complex organisms reassociates very slowly, a buffer which increases the reassociation rate was desirable. Therefore, the reassociation characteristics in several buffers were studied. Comparison was made between two distantly related organisms \([\text{the starry flounder (Platichthys stellatus)} \text{ and steelhead}]\), two closely related individuals \([\text{chinook (Oncorhynchus tshawytscha)} \text{ and steelhead}]\), and two distinctly different populations of the same species \(\text{(Salmo gairdneri)}\). The genomes of these animals were examined and compared in complexity, sequence similarity, and composition with respect to repeated sequences.
METHODS AND MATERIALS

DNA Preparation

DNA was obtained from *Escherichia coli B* and from mature fish testes. A starry flounder (*Platichthys stellatus*) from Yaquina Bay, Oregon, a chinook jack (*Oncorhynchus tshawytscha*) from the Siletz River, Oregon, and steelhead from the Siletz and Rogue Rivers in Oregon and from Sashin Creek in southeastern Alaska were used. The testes were ground in a blender with dry ice prior to extraction (see Appendix I for stepwise procedure).

The method of Marmur, modified by addition of 0.04% deoxycholate to the initial suspension, was employed [14, 17]. DNA preparations were stored at 4°C over chloroform. Carefully prepared DNA had a molecular weight of about $1.46 \times 10^7$ as determined by sedimentation ultracentrifugation.

For reassociation experiments, DNA dialyzed into the desired buffer was passed through a syringe several times and twice sheared by passage through a small orifice at a pressure drop of about 40,000 psi. The pressure was obtained from an American Instrument Company air-driven plunger pump, catalog number 46-13715. Chloroform extraction, centrifugation at 5-10,000 rpm for about 10 minutes, passage through a Metricel GA-6, 45μ pore size, cellulose-acetate filter [2] and dialysis prepared the DNA.
for reassociation. After centrifugation, a thick emulsion appeared at the interface of the two liquids; decanting and filtration separated the chloroform phase and the emulsion from the aqueous phase which contained the DNA. The nature of the emulsion was not investigated.

**Melt Curves**

A Gilford 2400 recording spectrophotometer with automatic reference compensator and thermally-controlled sample chamber, a Haake FJ constant temperature circulator with ethylene glycol, and teflon stoppered Perkin-Elmer 1 cm or Hellma 1 mm path length cells were used for both melt curves and reassociation experiments. DNA solutions were bubbled with helium, introduced into the cells, layered with mineral oil, and stoppered tightly just prior to use.

For the melt curves, the temperature of the sample was raised from ambient to approximately 100°C. The melting temperature (Tm) is the temperature at which the absorbance of the sample (at 260 nm) is midway between the minimum and maximum absorbances observed during the melt.

A reasonably constant temperature increase of the spectrophotometer sample compartment was obtained by disconnecting the thermostat of the circulator. Temperature increase was somewhat
rapid producing a lag between the cuvette temperature and the recorded temperature; the results, however, were consistent. After correction was made for thermal expansion of the buffer [13], the guanine plus cytosine content was calculated [15]; results for steelhead DNA were consistent with published values of about 43% [10] or a $T_m$ equal to 86.9°C in SSC ($0.15 \text{ M NaCl} + 0.015 \text{ M Na citrate, pH } 7$).

Melts were made at each stage of DNA extraction and criteria for DNA purity established. A hypochromicity (change in absorbance during the melt divided by the maximum absorbance value) greater than 0.25 was selected [6]. Where melts were not run, a ratio of absorbance at 260nm to absorbance at 280nm greater than 1.85 was adopted as a criterion for protein absence and a ratio of absorbance at 260nm to absorbance at 230nm greater than 2.3 for RNA absence [14].

**Sheared DNA Size**

Electron photomicrographs of sheared Sashin Creek steelhead DNA were made by Alfred Soeldner of Oregon State University's Botany Department on a Phillips Model 300 electron microscope. Grids were prepared from sheared DNA in $0.1 \text{ M ammonium acetate}$, and shadow cast with aluminum or carbon. The magnification factors of both the electron microscope and the photograph processing were employed in determining the DNA strand length from eight by ten inch prints. Single strand lengths of 380 to 540 nucleotides
were observed.

Alkaline cesium chloride sedimentation determinations were made with the help of Thomas Hinds of Oregon State University's Department of Biophysics and Biochemistry using a Spinco Model E ultracentrifuge equipped with a photoelectric scanner. Corrections for the cesium salt and solvent viscosity were taken from Bruner and Vinograd [7]. The sedimentation values measured were extrapolated to a solution density of one [7]. Wetmur and Davidson's [26] equation for converting alkaline single stranded DNA sedimentation coefficients to strand length in nucleotides was employed; values obtained were 380 and 450 nucleotides, corresponding well to the length obtained from the electron micrographs [26]. The lengths are small enough to avoid gel formations and probably yield reasonable kinetics [2].

Reassociation of DNA

The DNA was either dissociated by a slow, monitored melt prior to reassociation or by melting it quickly in heated NaCl. For the quick melt, the DNA solutions were bubbled with helium, placed in cuvettes, layered with mineral oil, and tightly stoppered. Cuvettes were placed in the spectrophotometer cuvette holder which accommodated four cells. One cell position was kept empty and used as a reference throughout the experiment. The holder, with a
thermometer inserted in the reference position, was wrapped in aluminum foil. The foil-wrapped holder was immersed in granular sodium chloride and gently heated over a burner. When the thermometer had registered over $100^\circ C$ for two minutes—usually as high as $110^\circ C$—the holder was removed, unwrapped, and quickly transferred to the chamber of the spectrophotometer, which was maintained at the temperature at which the reassociation was to be made (i.e., $60^\circ C$ or $70^\circ C$); recording was begun immediately.

Approximately 15 minutes were required for the temperature to equilibrate. This time had no significant effect on measurements made after long times. It was assumed that there was no viscosity increase from high concentrations of sheared DNA and that families of sequences which diverged sufficiently so as to yield no significant amount of reassociation under conditions employed are different [18]. No correction was made for the small initial absorbance decrease which accompanies the temperature drop initiating the reassociation.

Hydroxyapatite (HAP) chromatography was employed for fractionating DNA according to the amount of base-sequence repetition. Calbiochem Biogel HTP was used in a jacketed column. The resin support of the column was recessed to assure temperature control over the complete length of the HAP. Whatman GF/C glass filters (2.4 cm) were used over a perforated teflon disc support for the HAP. A Haake FJ constant temperature circulator provided
temperature control. Absorbance readings were made at both 260nm and 320nm to monitor HAP leakage [3]. Fractionation was done as described by Britten and Kohne [4].

Slowly reassociating DNA samples were lyophilized after fractionation to increase the DNA concentration for subsequent runs.

The second order nature of DNA reassociation was tested by reassociating different concentrations of steelhead DNA in 1 M NaClO₄ + 0.01 M Tris-HCl, pH 7.0 (ClO₄). The amounts of reassociation that had occurred at the same Cₒ t value in different runs were compared. Cₒ t is the product of the total DNA concentration and time, expressed as (moles of nucleotides)(seconds)(liters)⁻¹. The Cₒ t at which one half of the DNA has reassociated equals the reciprocal of the rate constant (see Appendix II, equation 2). Except for the amount of reassociation observed for high DNA concentrations (10-25 OD) at low Cₒ t's and that observed at low concentrations (1OD) at large Cₒ t's, the values observed were quite similar over Cₒ t values ranging from 0.1 to 200. Deviations observed with high and low concentrations probably result from the time required for initial temperature equilibration and from evaporation.
Calculation of Reassociation Kinetics

Most reassociation results are plotted as percent reassociation
\[ \frac{A(0) - A(t)}{A(0) - A(\infty)} \]
or as relative decrease in absorbance, \[ \frac{A(0) - A(t)}{A(0)} \]
\[ \log C \ t, \text{ where } A(0) \text{ is the maximum (initial) absorbance, } A(t) \text{ the absorbance at time } t, \text{ and } A(\infty) \text{ the absorbance at infinite time (total reassociation). } A(t) \text{ is used here rather than } C(t) \text{ because absorbance rather than concentration is measured in the experiments. The two values are not interchangeable (see Appendix II).}

Using an average molecular weight of 662 for a sodium nucleotide-pair and assuming that 33.4 µg/ml of dissociated DNA has an absorbance of one optical density unit (OD) [26], \[ 1 C_0 \]
equals 2.752 OD-hours, a somewhat higher value than that used by Britten and Kohne [5].

Native two-stranded DNA has a lower absorbance than completely reassociated DNA both because of strand scission caused by high temperature exposure and because of incomplete reassociation, a result of random reassociation of pairs not sheared at the same site or of repeated sequences interspersed among less redundant sequences that produces a product with "gaps".

An estimate of absorbance at infinite time may be obtained in two ways. A plot of A(t) or A(t)^{-1} against (A(0)t)^{-1} may be extrapolated to (A(0)t)^{-1} equals zero to obtain an estimate of A(t) at
infinite time. Alternatively, for organisms with nonrepeated DNA, one may utilize the linearity of a plot of \( \frac{A(0) - A(\infty)}{A(t) - A(\infty)} \) against time, a property of second order kinetics [23]. (See Appendix II)

An iteration procedure whereby the value of \( A(\infty) \) which yielded the minimum regression error was used to estimate \( A(\infty) \). Relative \( A(\infty) \) values obtained for \textit{E. coli}, harvested in the stationary phase, were assumed to equal those of other DNA's studied. Values obtained from \textit{E. coli} DNA in different buffers were similar (see Table 1).

A second method of plotting reassociation data was used to illustrate the relative size of repetitious components within the genome of an organism. Since the rate at which DNA reassociates at a given \( C_{ot} \) is proportional to the concentration of sequences reassociating, the slope derived from the reassociation curve (percent reassociation versus \( \log C_{ot} \) plotted against \( \log C_{ot} \)), will indicate the relative amount of DNA reassociating at each \( C_{ot} \). This plot may be more conveniently read as \((-\log_{10} \text{repetition frequency}) \) versus the relative quantity of DNA, where \((-\log_{10} \text{repetition frequency}) \) is zero under the peak of unique DNA.
Table 1. Estimates of maximum relative decrease in absorbance of sheared *E. coli* DNA from the best least squares fit to the straight line $\frac{A(0)-A(t)}{A(t)-A(\infty)}$ on time (t) obtained by varying $A(\infty)$. The number of runs comprising each value is in parentheses. The buffers are defined in the text.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Buffer</th>
<th>SSC</th>
<th>PB</th>
<th>2SSC</th>
<th>CI04</th>
<th>HC104</th>
</tr>
</thead>
<tbody>
<tr>
<td>60°C</td>
<td></td>
<td>0.2137(8)</td>
<td>0.2147(6)</td>
<td>0.2305(5)</td>
<td>0.2220(4)</td>
<td>0.2217(5)</td>
</tr>
<tr>
<td>70°C</td>
<td></td>
<td>0.2175(2)</td>
<td>0.2091</td>
<td>0.2311</td>
<td>0.2099\textsuperscript{a}</td>
<td>0.2172</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Influenced by evaporation.
RESULTS AND DISCUSSION

Buffer Comparisons

The buffers investigated were SSC* (.15 M NaCl + .015 M Na citrate, pH 7.0), 2SSC (.30 M NaCl + .030 M Na citrate, pH 7.0), PB (.12 M PO$_4$ equimolar in mono- and di-basic Na-PO$_4$), ClO$_4$(1.0 M NaClO$_4$ + .01 M tris-HCl, pH 7.1), and HClO$_4$ (.5 M NaClO$_4$ + .01 M tris-HCl, pH 7.1). Typical curves of E. coli B DNA sheared in SSC and reassociated in these buffers are shown in Figure 1.

The reassociations of E. coli DNA were plotted as percent reassociation on log C t. The curves were characterized by the C t intercept at zero reassociation and the slope of the straight line portion of the latter part of the reassociation curve. The intercept is proportional to the reciprocal of the reassociation rate constant if one assumes the slopes of the buffers are equal. The values obtained from these lines were used for comparisons among buffers, between reassociations at 60°C and 70°C in the same buffer, between E. coli DNA sheared in the buffer in which it was reassociated and E. coli DNA sheared in SSC, and between DNA slow melted and DNA quick melted prior to reassociation (see Table 2).

Intercept values of DNA reassociation curves indicate that

* Standard Saline-Citrate.
Figure 1. Reassociation of \textit{E. coli} DNA sheared as described in the text in the buffers,

- SSC
- 2SSC
- PB
- HClO4
- ClO4
Table 2. Properties of the straight line portion of the plot of percent reassociation on log C \(_t\) of E. coli DNA reassociated in different buffers and at different temperatures. Numbers in parentheses indicate the number of runs comprising each value. Underlined values indicate a significant departure from the value obtained from runs in the same buffer but sheared in SSC and reassociated at 60°C. The buffers are described in the text.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Buffer</th>
<th>SSC</th>
<th>PB</th>
<th>HC1O4</th>
<th>ClO4</th>
<th>2SSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>[cation]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interception of PB(60°C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60°C Sheared in SSC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C (_t) 1/2</td>
<td>2.4</td>
<td>3.3</td>
<td>.53</td>
<td>.36</td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>OD range</td>
<td>1.2-2.1</td>
<td>1.7-2.2</td>
<td>1.1-2.1</td>
<td>1.2-2.1</td>
<td>1.6-2.1</td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>10.8</td>
<td>10.5</td>
<td>11.5</td>
<td>11.4</td>
<td>12.4</td>
<td></td>
</tr>
<tr>
<td>B.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interception of PB(60°C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70°C Sheared in SSC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C (_t) 1/2</td>
<td>3.0</td>
<td>4.0</td>
<td>.56</td>
<td>.45</td>
<td>.87</td>
<td></td>
</tr>
<tr>
<td>OD range</td>
<td>1.8</td>
<td>1.8</td>
<td>1.5</td>
<td>1.5</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>10.1</td>
<td>9.84</td>
<td>9.75</td>
<td>9.05</td>
<td>10.0</td>
<td></td>
</tr>
</tbody>
</table>

Continued
Table 2--Continued.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Buffer [cation]</th>
<th>SSC</th>
<th>PB</th>
<th>HC1O4</th>
<th>ClO4</th>
<th>2SSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.</td>
<td>Intercept</td>
<td>.203(3)</td>
<td>.246(2)</td>
<td>.0411(2)</td>
<td>.0369(1)</td>
<td>.0509(1)</td>
</tr>
<tr>
<td>60°C Sheared and reassociated in the same buffer</td>
<td>Intercept</td>
<td>1.23</td>
<td>.861</td>
<td>6.06</td>
<td>6.75</td>
<td>4.90</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>10.6</td>
<td>11.6</td>
<td>11.4</td>
<td>10.6</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>OD range</td>
<td>1.1-1.3</td>
<td>1.3</td>
<td>.8-1.3</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>D.</td>
<td>Intercept</td>
<td>.176(1)</td>
<td>.163(1)</td>
<td></td>
<td></td>
<td>.0381(1)</td>
</tr>
<tr>
<td>60°C Sheared and reassociated in the same buffer</td>
<td>Intercept</td>
<td>1.42</td>
<td>1.54</td>
<td></td>
<td></td>
<td>6.54</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>10.7</td>
<td>9.70</td>
<td></td>
<td></td>
<td>10.8</td>
</tr>
</tbody>
</table>
the buffer in which the DNA is sheared significantly effects the reassociation rate (compare Table 2A with 2D). The rate at which DNA reassociated decreased when the DNA was slow melted rather than quick melted prior to reassociation; compare values of intercepts in Table 2C and 2D. The differences observed from the two melting procedures may result from the longer time needed for the temperature to equilibrate at the reassociation temperature after dissociation. It would seem, however, that over long time periods this initial delay would be insignificant. Comparison of Tables 2A and 2B shows that the increase in temperature from 60°C to 70°C retards the reassociation.

The slopes of the curves, although similar, did have some significant differences. That of DNA reassociated with 2SSC was slightly higher than those of the others; temperature increase caused a decrease in the slopes.

ClO4 increases the reassociation rate about 13 times that of PB which is used in HAP chromatography and nearly nine times that of SSC. The large rate increase produced by the use of ClO4 makes it quite valuable for work with eucaryotes whose large genome size necessitates long reassociation. Similarly, PB or another low cation concentration buffer could be conveniently used to study faster reassociating DNA. Knowledge of the relative rates of reassociation allows comparison of experiments run in different
buffers. Conditions selected for racial or interspecies comparisons of DNA sheared in SSC were ClO₄ at either 60°C or 70°C.

The rate constants obtained from C₀t 1/2 values were somewhat slower than some published values, a result of shorter strand length for buffers of comparable cation concentration (see Table 3).

**The Steelhead Genome**

A steelhead DNA reassociation curve, a composite of three separate runs, and a rate of reassociation versus log C₀t curve derived from the reassociation curve are shown in Figures 2 and 3, respectively. The reassociation curve shows an initial drop in relative absorbance followed by a slow, but continuous, decrease over the rest of the experiment. That there is no definite final drop may be an indication either that large C₀t values are obscured by strand scission or evaporation or that the unique region has not been reached. From both this reassociation curve and the reassociation curve of large C₀t DNA separated from the bulk DNA (Figure 8), the half C₀t of reassociation in ClO₄ at 70°C may be estimated to be on the order of 200 to 800. This C₀t range indicates that a haploid genome of nonrepeated steelhead DNA is about 2 to 8 x 10⁹ nucleotide pairs [5], if one assumes a half C₀t of 0.45 (see Table 2) for *E. coli* whose genome size is 4.5 x 10⁶ nucleotide pairs [8]. The value for steelhead is approximately that of calf thymus DNA estimated by
Figure 2. The reassociation of steelhead DNA. The reassociation was done in CI04 using DNA sheared twice at 40,000 psi. The curve is a composite of three separate runs, the initial concentrations of which were 1, 10, and 25 OD.

Figure 3. The rate of reassociation of steelhead DNA (Figure 2) 
\[ \frac{\Delta A_{o} - A_{t}}{A_{o}} \] 
\[ \frac{\Delta \log_{10} C_{t}}{C_{o}} \] plotted against log C o. This curve may be interpreted as the relative amount of DNA plotted against (-log repetition frequency). The percentages on the graph indicate the extent of completion of the reassociation.
Table 3. Second order rate constants ( liter mole−1 sec−1) for the reassociation of E. coli DNA. Buffers are defined in the text.

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH 13</th>
<th>Reassociation temperature</th>
<th>Rate constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4 M NaCl pH 7.0</td>
<td>30</td>
<td>70°C</td>
<td>1.4 [15]</td>
</tr>
<tr>
<td>1.9 SSC</td>
<td>33</td>
<td>68°C</td>
<td>2.37 [16]</td>
</tr>
<tr>
<td>2SSC</td>
<td>6</td>
<td>70°C</td>
<td>1.15</td>
</tr>
<tr>
<td>SSC</td>
<td>6</td>
<td>60°C</td>
<td>0.417</td>
</tr>
<tr>
<td>ClO4</td>
<td>6</td>
<td>70°C</td>
<td>2.22</td>
</tr>
</tbody>
</table>

Britten and Kohne [5]. This measurement indicates the maximum amount of information (complexity) an organism has coded in its base sequences.

As seen in Figure 3, the initial drop in relative absorbance results from a highly repeated (5 x 10^3 - 10^4 copies) portion, while the long slow decline indicates the presence of DNA at many frequencies of repetition intermediate (1-10^3 copies) to the highly repeated DNA and the unique DNA. As much as 60-80% of the genome is repeated to some extent; the uncertainty in this value results from failure to definitely identify the unique area. Although time required initially to lower the temperature of the melted DNA to the temperature of reassociation makes early reassociation values
unreliable, one may estimate (Figure 3) that as much as 25% of the genome is highly repeated.

**Comparison of Steelhead and Starry Flounder DNA**

The starry flounder was selected for preliminary comparison to the steelhead both because of its availability and its distant common ancestry. Although one would expect some similarity in base sequences common to all vertebrates, sufficient divergence should have occurred so that one might expect resolvable differences. The resolution achieved by the comparison of reassociation curves of the DNA's of the two organisms indicated that the technique was sufficiently discriminating for use on more closely related organisms.

The reassociation shown in Figure 4 demonstrated large differences between the genomes of the two organisms. The flounder DNA reassociation curve shows a genome comprised predominantly of two components, a fast fraction somewhat smaller than that of the steelhead—one third of the DNA—and a very large nonrepeated portion representing two thirds of the genome (see Figure 5). Unlike the steelhead, the flounder has no significant amount of DNA repeated at frequencies intermediate to the very rapid and slowly reassociating components. That the fast fractions of both organisms appear to occur at the same C t may be misleading in that the
Figure 4. The reassociation of the DNA's of starry flounder — ••, initial concentration of 18.4 OD, Rogue Steelhead — ••, initial concentration of 18.6 OD, and a mixture of equal parts of both — — , initial concentration of 18.5 OD — •• represents a theoretical curve for the mixture of flounder and steelhead DNA assuming that no sequence similarities exist. The reassociation was run at 60°C in ClO4 with DNA sheared as described in the text. The runs were made in 1 mm cuvettes.

Figure 5. The slope of the reassociation curves shown in Figure 4 plotted against log C t, starry flounder — ••, steelhead — ••, and an equal parts mixture of the two — — . This curve may be read as relative amount of DNA versus (-log repetition frequency). The percentages on the graph indicate the extent of completion of reassociation.
The genome size of the nonrepeated DNA of the flounder
which has a temperature corrected half \( C_0 \) of approximately 110
is considerably smaller than that of the steelhead, whose nonrepeated
genome has a half \( C_0 \) of 200 or greater.

The theoretical curve expected from a mixture of steelhead
and flounder DNA (Figure 4) assumes no similarities in sequences
and was derived from the reassociation curves of unmixed DNA's
of the two organisms run at the same time as the mixture. The
curve was plotted as one half the sum of the relative absorbancies
of the two unmixed DNA's at the same \( C_0 \) with twice that \( C_0 \) (it
takes twice as long for a one in two diluted sample to reassociate
to the same relative extent as the undiluted sample). Sequence
similarities between the two organisms would appear as more
rapid decreases in the relative absorbance of the mixture than are
predicted by the theoretical curve. The faster decrease results
from reassociation of sequences that are common to both organisms
and are, therefore, more repeated than the theoretical curve
predicts.

The theoretical and experimental curves actually coincide
during the longer \( C_0 \) 's of reassociation, an indication of nearly
total dissimilarity of the less repeated sequences of the two species.
The differences between the two curves at small $C_0$'s are probably due more to the time involved in initial temperature equilibration than to sequence similarities. With the possible exception of very early or very late $C_0$'s, the steelhead and flounder appear to be quite dissimilar in sequencing, genome structure, and genome size. The DNA's of these organisms have diverged to such a degree that there is no measurable similarity between them using this comparison technique. The large differences in genome structure indicate that base changes alone do not account for the amount of divergence observed between the two species. Such great divergence between the two organisms could result from extensive duplication and deletion and possibly from a process such as saltatory replication which is described by Britten and Kohne as the incorporation into the genome of a short sequence of DNA replicated a large number of times [5].

**Chinook-Steelhead Comparisons**

Comparisons between the DNA's of two closely related species, chinook and steelhead, were made to further test the discriminating ability of the technique described in the previous section. The reassociation curves (Figure 6), representative of 15 comparison experiments, show that the genome of the chinook jack is quite similar to that of the steelhead. The initial rate of reassociation
Figure 6. DNA reassociation curves of chinook jack — — at 25.3 OD, Siletz steelhead — — at 25.7 OD, and an equal parts mixture of the two — — at 26.2 OD. The reassociations were run at 60°C in ClO4 in 1 mm cuvettes with DNA sheared as described in the text.

Figure 7. DNA reassociation curves of HAP fractionated fast reassociating DNA from chinook jack — — at 1.34 OD, steelhead — — at 1.33 OD, and an equal mixture of the two at 1.32 OD. The reassociations were run at 70°C in ClO4 in 1 cm cuvettes. The DNA was sheared prior to HAP fractionation.

Figure 8. DNA reassociation curves of HAP fractionated slowly reassociating DNA from chinook jack — — at 8.25 OD, steelhead — — at 7.04 OD, and a mixture of equal amounts of the two — — at 7.78 OD. The reassociation was run at 70°C in ClO4 in 1 mm cuvettes. The DNA was sheared prior to HAP fractionation.
of chinook DNA appears to be slightly faster than that of the steelhead, implying that the chinook genome has both a slightly larger and a somewhat more repeated quickly reassociating fraction. Because the two curves are so similar in shape and the early reassociation is influenced by the temperature equilibration process, the early differences between the curves may not be significant. The major difference between these species observed in all experiments was the slower rate of reassociation, relative to those of the chinook and steelhead references, of the less repeated sequences of the chinook-steelhead DNA mixture. This slower rate indicates some differences in base sequencing between the two organisms. A similar number of common or similar sequences exist, however, because complete dissimilarity would result in a mixture curve shaped the same as the steelhead or chinook reassociation curves, but displaced a $C_o t$ of two greater than the present $C_o t$ plots--it would reassociate more slowly by a factor of two.

Attempts were made to amplify the observed differences between the chinook and steelhead. The DNA of both organisms was fractionated on HAP to obtain very fast and slowly reassociating components. Samples were lyophilized when more concentrated solutions were required. The very fast fractions were compared (Figure 7) by the technique previously described, nearly complete reassociation of both samples and the mixture occurred before the
equilibrated. The half \( C_{0.03} \) of these reactions, a maximum estimate, was approximately 0.03, a value which corresponds to an effective genome size of 300,000 nucleotide pairs. That the DNA's did not reassociate to the extent the theoretical value (Table 1) predicts results primarily from incomplete separation of this fast fraction from slower ones. The DNA mixture follows the two references closely indicating either that there is great similarity between the two organisms in this fraction, or that the genome size estimate has been grossly overestimated because of the slow initial temperature equilibration.

Comparisons of the slowly reassociating DNA's of chinook and steelhead (Figure 8) again displayed the sequence differences observed in the unfractionated DNA comparisons. The slope of the straight line portion of the curve is somewhat less than one would expect from single copy DNA sequences suggesting the presence of DNA with very low repetition frequencies with the single copy DNA.

That the two species are distinguishable from differences in base sequencing but not in genome structure (degree of repeated-ness) is an indication that base changes are the most significant process in the recent divergence of these two organisms.
Comparisons of the DNA's of Different Steelhead

Having demonstrated the ability to differentiate between two closely related species, the comparison technique was employed to determine the possibility of differentiating between populations of steelhead. Comparisons were made between Siletz and Rogue River steelhead and between Rogue River and Sashin Creek steelhead. The long distance separating the Rogue and Sashin Creek and the difference in timing of the spawning runs virtually precludes the possibility of direct inter-breeding. Individuals from these localities may surely be considered as representatives of discrete breeding "groups".

No significant or repeatable differences were found in the comparisons either at 60° or 70°C. A typical comparison is shown in Figure 9.

That the possible differences between these populations is unmeasurable may be due to poor resolution of the technique. It is also possible that the time since the separation of the two groups was so short that no significant amount of divergence has occurred. Lastly, the localities studied may in fact contain representatives of a single population.
Figure 9. DNA reassociation from Sashin Creek steelhead at 21.8 OD, Rogue River steelhead at 19.5 OD, and a mixture of equal amounts of the two at 20.6 OD. The reassociations were made at 60°C in ClO4 in 1 mm cuvettes using DNA sheared as described in the text.
Perspectives

The long times required by each reassociation experiment restricted the study to a small number of species. The number of individuals of each species used in this study was quite small. The observation of differences between two such closely related species as chinook and steelhead shows that relatively small differences can be resolved. If any significant differences do exist between the two populations of steelhead studied, the differences should be discernible. In addition, when one makes a bulk comparison of the DNA's, one would expect many fixed (and thus identical) loci and many different only by a few base changes. Because of the large number of loci in the genome the probability of an extreme departure (measured in base changes) from the mean is quite small. It seems likely, therefore, that numerous individuals of a species will not represent that species more accurately than will a few.

The time required at the initiation of reassociation experiments for temperature equilibration makes the study of quickly reassociating components difficult. It is possible that the DNA extraction procedure selected against some base sequences. The DNA extracted in this case would less accurately reflect the genome of the organism. More information concerning the quickly reassociating DNA may be obtained by fractionation followed by reassociation
in a buffer that retards reassociation or by reassociation at a higher temperature.

The most interesting observation made in this study is the total dissimilarity between the DNA's of starry flounder and steelhead. While the small number of species studied prevents one from speculating wildly, this divergence within the fishes suggests that the investigation of the DNA's of more species may provide additional surprises of instructive nature.
LITERATURE CITED


APPENDIX I

PROCEDURE FOR DNA EXTRACTION [14, 17]

1. Excise maturing testes and place on dry ice; store frozen.
2. Cool blender until well frosted by blending dry ice; shatter frozen testis by striking, then blend until pulverized; store frozen.
3. Add (very roughly) 10 g of pulverized testis to 125 ml of .15 M NaCl + .1 M EDTA + .04% deoxycholate, pH 8, add 10 ml 25% sodium lauryl sulphate (SLS), and blend in a room temperature blender a very short time (long enough to thoroughly mix the components).
4. Heat the mixture to 60°C for 10 minutes.
5a. Add 5 M NaClO₄ to produce a solution 1 M in NaClO₄ and shake well with an equal volume of 24:1 chloroform:isoamyl-alcohol for approximately 30 minutes. OR
5b. Shake with an equal volume of 90% phenol.
6. Separate with phases by centrifugation at 5-10,000 rpm for 5-10 minutes; retain the aqueous phase.
7. Repeat 5 and 6, or go to 8.
8. Precipitate the DNA from the aqueous phase by adding two volumes of cold 95% ethanol; spool the DNA on a glass rod.
9. Dissolve the DNA in .1XSSC. (SSC is .15 M NaCl + .015 M Na citrate, pH 7).

10. Add 10XSSC to attain SSC.

11. Repeat 5 to 10 until little protein is left at the interface after centrifugation, usually 2-3 times.

12. Add RNase (50 μg/ml, .2% in .15 M NaCl, pH 5; and heated to 80°C for ten minutes to deactivate DNase) and incubate for about 30 minutes at 37°C.

13. Add pronase or protease from *Streptomyces griseus* (type V) (50 μg/ml self digested at 37°C for two hours) for two hours at 37°C.

14. Make the solution .5% in SLS; repeat 5 to 10 until no protein is left at the interface after centrifugation.

15. Precipitate the DNA with two volumes of ethanol and spool it on a glass rod; dissolve it in 13.5 ml of .1xSSC, add, while stirring constantly, 15 ml of 3.0 M Na acetate + .001 M EDTA, pH 7; add 15 ml of cold isopropanol dropwise while spooling the precipitating DNA on a glass rod.

16. Stir the spooled DNA in progressively increasing concentrations of ethanol (70%-95%).

17. Repeat 16, usually once.
18. Dissolve the DNA in the desired buffer—usually SSC—and dialyze extensively.
APPENDIX II

SECOND ORDER EQUATIONS

A second order process is defined by the following differential equation:

\[ \frac{dc}{dt} = -k C^2. \] (1)

Where \( C(t) \) is the concentration of unreassociated DNA, the equation may be solved to yield:

\[ C(t) = \left( k t + \frac{1}{C(0)} \right)^{-1} \quad \text{or} \quad \frac{C(0) - C(t)}{C(t)} = C(0) k t. \] (2)

When hypochromicity is used rather than concentration, \( C(t) \) must be replaced by \( A(t) - A(\infty) \), which indicates the amount of DNA not yet reacted, the differential equation is:

\[ \frac{dA(t)}{dt} = -(A(t) - A(\infty))^2 k, \] (3)

which may be solved to yield:

\[ \frac{A(0) - A(t)}{A(t) - A(\infty)} = \left( \frac{A(0) - A(\infty)}{A(0)} \right) A(0) k t. \] (4)

Since the maximum hypochromicity (\( H \)) is equal to \( \frac{A(0) - A(\infty)}{A(0)} \), this solution may be written as:

\[ \frac{A(0) - A(t)}{A(t) - A(\infty)} = HA(0) k t. \] (5)
An alternative expression is:

\[
\frac{A(0)}{A(0) - A(t)} = \frac{1}{A(0)k} tH^2 + \frac{1}{H}.
\] (6)

Plots of \(\frac{A(0)}{A(0) - A(t)}\) versus \((A(0)t)\) provide estimates of both \(H\) and \(k\). If one assumes a value for \(H\) (see Table 1), the slope of \(\frac{A(0) - A(t)}{A(t) - A(\infty)}\) versus \(A(0)t\) estimates \(k\) directly.
If one assumes that the DNA of a reassociation experiment is comprised of two independently reassociating components ($C_1$ and $C_2$, $C_1 + C_2 = C$), each of which follows second order kinetics ($\frac{dC_1}{dt} = -k_1 C_1^2$, $\frac{dC_2}{dt} = -k_2 C_2^2$), the resulting reassociation will follow the equation:

$$C(t) = \left(k_1 t + \frac{1}{C_1(0)}\right)^{-1} + \left(k_2 t + \frac{1}{C_2(0)}\right)^{-1}.$$

Examples of second order curves where $k$ has different values, as well as mixed component second order curves are shown in Figure 10.
Figure 10. Theoretical second order curves for different rate constants \( k \), 1.0, 0.1, .01, and .001; and two theoretical two-component curves, one of whose rate constants are .01 and 1 (the two components are present in equal amounts) and the other whose rate constants are .01 and 1.