The genome composition and DNA base sequence homologies of several diverse species of fish were examined. The fishes were characterized from spectrophotometric studies of reassociation and hybridization of their DNAs and from the thermal stabilities of hybrid duplexes.

Spectrophotometric studies indicated that repeated sequences comprise 35-40% of the genomes of starry flounder (*Platichthys stellatus*), ling cod (*Ophiodon elongatus*), and buffalo sculpin (*Enophrys bison*), and approximately 60% of the genomes of the salmonids (*Salmo gairdneri*, *Oncorhynchus tshawytscha*, *Prosopium williamsoni*, and *Thymallus arcticus*) and Pacific lamprey (*Entosphenus tridentatus*). The information content was relatively high in the steelhead (*Salmo gairdneri*), $1.6 \times 10^9$ base pairs, and lamprey, $2.3 \times 10^9$ base pairs,
as compared to $0.75 \times 10^9$ in starry flounder. Hybridization studies indicate that closely related species share substantial base sequence similarities while distantly related fish share few or none. Hydroxyapatite chromatography was used for examining the thermal stabilities of DNA hybrids. This technique provided a means for separating closely related species, such as the salmonids. It was concluded that the specialization in fishes is accompanied by a reduction in information content and a decrease in repeatedness in their genomes. These observations are consistent with the loss of structural components and decrease in plasticity generally accompanying specialization.

Quantitative differences were found in the base sequences of closely related species and, although differences between populations and races were not resolved, the data suggest that with sufficient refinement of the methodology, such separations might be made.
Comparison of the Reassociation and Hybridization Properties of the DNAs of Several Species of Fish

by

Anthony John Gharrett

A THESIS submitted to Oregon State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June 1975
ACKNOWLEDGEMENTS

I am indebted to my committee, Drs. George S. Beaudreau, Peter S. Dawson, and Kenneth E. Rowe, for their support, interest, and help. Interest in this project was stimulated by Dr. Raymond C. Simon, my major professor, and I greatly appreciate his non-interference policy. Special thanks go to Dr. John D. McIntyre who provided help and incentive, both financial and moral.

Specimens were provided by Mr. Morris Barker, Mr. Don Ratliff of Portland Gas and Electric Company, the Oregon Wildlife Commission, and the Oregon Fish Commission.

The labeled DNA was sonicated by Mr. Gary Thorgard of the Department of Genetics at the University of Washington. Dr. Thomas Hinds of Oregon State University's Department of Biochemistry and Biophysics provided the necessary expertise for ultracentrifugation.

This investigation was supported by the Oregon Cooperative Fishery Unit; cooperators are the U.S. Fish and Wildlife Service, the Oregon Wildlife Commission, the Oregon Fish Commission, and Oregon State University.
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I  INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II MATERIALS AND METHODS</td>
<td>6</td>
</tr>
<tr>
<td>Buffers</td>
<td>6</td>
</tr>
<tr>
<td>DNA Preparation</td>
<td>6</td>
</tr>
<tr>
<td>Sheared DNA Size</td>
<td>8</td>
</tr>
<tr>
<td>Melt Curves and DNA Purity Criterion</td>
<td>9</td>
</tr>
<tr>
<td>Reassociation of DNA and Hydroxyapatite</td>
<td>11</td>
</tr>
<tr>
<td>Fractionation</td>
<td>11</td>
</tr>
<tr>
<td>Thermal Stability Elutions</td>
<td>13</td>
</tr>
<tr>
<td>Calculation of Reassociation Kinetics</td>
<td>14</td>
</tr>
<tr>
<td>III RESULTS</td>
<td>17</td>
</tr>
<tr>
<td>Optical Comparisons of Reassociation and Hybridization Properties</td>
<td>17</td>
</tr>
<tr>
<td>Optical Comparisons of Closely Related Species</td>
<td>30</td>
</tr>
<tr>
<td>Melting Profiles of DNA Hybrids</td>
<td>36</td>
</tr>
<tr>
<td>IV DISCUSSION</td>
<td>43</td>
</tr>
<tr>
<td>V PERSPECTIVES</td>
<td>48</td>
</tr>
<tr>
<td>REFERENCES CITED</td>
<td>50</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>54</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 <em>E. coli</em> 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 $A(\infty)$.</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>The size of the fast component and the minimum complexity of the DNAs of ten species of fish and the extent of similarity between them.</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>Properties of the straight line portion of the plot of percent reassociation on log C_{ot} of <em>E. coli</em> DNA reassocaited in different buffers and at different temperatures.</td>
<td>64</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>The reassociations of the DNAs of ten species of fish.</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>The reassociations of the DNAs of Deschutes rainbow, Deschutes steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>The reassociations of the DNAs of Deschutes steelhead, starry flounder, and a mixture of equal parts of the two DNAs.</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>The reassociations of the DNAs of buffalo sculpin, ling cod, and a mixture of equal parts of the two DNAs.</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>DNA reassociation curves of chinook jack, Siletz steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>DNA reassociation curves of HAP fractionated fast reassociating DNA from chinook, steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>DNA reassociation curves of HAP fractionated slowly reassociating DNA from chinook, steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>32</td>
</tr>
<tr>
<td>8</td>
<td>DNA reassociation curves of Deschutes rainbow, Deschutes steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>34</td>
</tr>
<tr>
<td>9</td>
<td>DNA reassociation curves of HAP fractionated fast reassociating DNA from Deschutes rainbow, Deschutes steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>DNA reassociation curves of HAP fractionated slowly reassociating DNA from Deschutes rainbow, Deschutes steelhead, and a mixture of equal parts of the two DNAs.</td>
<td>34</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>11</td>
<td>Melting profiles of duplexes produced by incubation of a mixture of unlabeled Deschutes steelhead and labeled Deschutes steelhead DNAs.</td>
<td>40</td>
</tr>
<tr>
<td>12</td>
<td>Melting profiles of duplexes produced by incubation of a mixture of unlabeled Deschutes rainbow and labeled Deschutes steelhead DNAs.</td>
<td>40</td>
</tr>
<tr>
<td>13</td>
<td>Melting profiles of duplexes produced by incubation of a mixture of unlabeled Sashin Creek steelhead and labeled Deschutes steelhead DNAs.</td>
<td>42</td>
</tr>
<tr>
<td>14</td>
<td>Melting profiles of duplexes produced by incubation of a mixture of unlabeled golden trout and labeled Deschutes steelhead DNAs.</td>
<td>42</td>
</tr>
<tr>
<td>15</td>
<td>Melting profiles of duplexes produced by incubation of a mixture of unlabeled dolly varden and labeled Deschutes steelhead DNAs.</td>
<td>42</td>
</tr>
<tr>
<td>16</td>
<td>Theoretical second order curves for different rate constants (k), and two theoretical two-component curves.</td>
<td>59</td>
</tr>
<tr>
<td>17</td>
<td>Reassociation of sheared \textit{E. coli} DNA in different buffers.</td>
<td>62</td>
</tr>
</tbody>
</table>
COMPARISON OF THE REASSOCIATION AND HYBRIDIZATION PROPERTIES OF THE DNAs OF SEVERAL SPECIES OF FISH

I. INTRODUCTION

The classification systems that have been developed for use on fishes are based, primarily, on morphological (phenotypic) characteristics. Such characters result from environmental modulation of genetic information and reflect varying degrees of environmental influence. The various criteria developed from the morphological characters of fish often result in diverse interpretations of the relationships between fish.

Because the genetic material (DNA) contains the information necessary for all structures, functions, and behaviors of an organism and is inherited, comparisons of DNA base sequences provide definite criteria for establishing phylogenetic relationships between organisms. This basis for comparison has been used for studies of bacteria (28) and T-series bacteriophage (29), and subsequently for primates (18) and rodents (20).

The objectives of this study were to develop methodology for the comparison of DNA base sequences in fishes and to determine the kinds and amounts of differences for a few diversely related fish. Since the study of the genetic material requires the use of the characteristics of DNA, a brief description of the biological and chemical
properties follows below.

Structurally, DNA consists of two helically wound polynucleotide chains. Each chain is comprised of the deoxyribose nucleotides: deoxyadenylic acid, thymidylic acid, deoxycytidylic acid, and deoxyguanylic acid. The two chains are held together by the hydrogen bonding of specific pairs of the deoxyribose nucleotides; hence, the sequence of one strand determines that of the other. The many possible linear arrangements of these deoxyribose nucleotides provides the means by which information is coded into the DNA; the order is called the base sequence. The total information content, the non-repeated sequences, is the complexity and in many organisms exceeds $10^9$ base pairs.

The DNA of an individual is inherited from the ancestors of that individual. Changes in base sequences (additions, deletions, or base alterations) at each generation occur at a very low rate. These changes and the rearrangement of information present are the basis for evolutionary processes. There is a very low probability that two identical or nearly identical sequences will occur by chance in unrelated individuals. For example, there are more than $10^{60}$ different possible sequences for a polynucleotide of 100 bases. Hence, the more similarities there are in the base sequences of two individuals, the more recent one would expect that their common ancestor existed.
A detailed study of the structure and function of the DNA of an organism would ultimately provide complete biological knowledge about that organism. Unfortunately the quantity of information is enormous and the mechanism for expressing it complex. The slow rate at which changes in the DNA occur, the low probability of chance similarity, and the existence of techniques which demonstrate similarities and differences between sequences from diverse organisms do, however, provide a basis for determining the evolutionary distance between the two organisms (15, 19).

Double stranded DNA may be dissociated into single strands and, under appropriate conditions, reassociated to form duplexes possessing much of the high specificity of base pairing occurring in the original-'native'-duplex (24, 28). This reassociation process follows second order kinetics (33, 37).

A second order process requires a collision of two components to form the product; hence, the rate of reaction is dependent upon the concentration of both components (see Appendix II).

Reassociated duplexes comprised of DNA strands from two different individuals—possibly different species or populations—occur when sufficient complementarity between heterologous strands exists. These duplexes may be defined as "hybrids". "Hybrid" formation enables one to estimate base sequence similarity of the DNAs from two sources (19), while the lower thermal stability of "hybrids", a
result of some mispairing, permits the comparison of very similar (but not duplicate) sequences.

The progress of reassociation may be measured most practically by observing the decrease in absorbance of ultraviolet light (260 nm) accompanying duplex formation or by the separation of double from single strands on hydroxyapatite (1). The former method measures the base pairing of reassociation and provides a convenient means for continual observation of a long term reassociation experiment. The latter method measures strand pairing and is useful for separating quickly reassociating DNA from slowly reassociating DNA, for following the progress of reassociation, and for examining the thermal stability of duplexes. Excellent review articles considering both the theory and methodology of reassociation and hybridization are available (16, 17, 36).

All organisms are typified by a genome, which is defined as the complement of nucleotide base sequences contained in a haploid chromosome set. Higher organisms (fishes included) possess some sequences repeated many times relative to other sequences in their genomes. Within the same genome, a gradation in the extent to which different sequences are repeated, approaching an excess of one million copies in the extreme, may exist (2, 4, 5).

The second order nature of the DNA reassociation process indicates that the repeated sequences will reassociate faster than the
nonrepeated sequences, since the more abundant sequences have a higher probability of collision. Analyses of the reassociation curves would, therefore, provide information on the relative amounts of repetition present within the genome (5).
II. MATERIALS AND METHODS

Buffers

The following is a list of buffers used in this study, the letter by which it is designated in the text, the abbreviation (if any) commonly used in the literature, and the composition:

A  SSC  0.15 M NaCl + 0.015 M Na-Citrate, pH 7.0

B  2XSSC 0.30 M NaCl + 0.03 M Na-Citrate, pH 7.0

C  PB  0.5 M NaH$_2$PO$_4$ + 0.5 Na$_2$HPO$_4$

D  0.12 PB  0.06 M NaH$_2$PO$_4$ + 0.06 M Na$_2$HPO$_4$

E  0.4PB  0.2 M NaH$_2$PO$_4$ + 0.2 M Na$_2$HPO$_4$

F  1.0 M NaClO$_4$ + 0.01 M tris-HCl, pH 7.1

G  0.5 M NaClO$_4$ + 0.005 tris-HCl, pH 7.1

DNA Preparation

Unlabeled DNA was prepared from mature fish testes which have high DNA content. The fish studied were starry flounder (Platichthys stellatus), ling cod (Ophiodon elongatus), and buffalo sculpin (Enophrys bison) from Yaquina Bay, Oregon, Carp (Cyprinus carpio) and shad (Alosa sapidissima) from Bonneville Dam on the Columbia River in Oregon, grayling (Thymallus arcticus) from the U.S. Bureau of Sport Fisheries and Wildlife Hatchery in Bozeman, Montana, whitefish (Prosopium williamsoni), steelhead and rainbow trout (Salmo
gairdneri) from the Deschutes River in Oregon, chinook salmon 
(Oncorhynchus tshawytscha) from the Siletz River in Oregon, Pacific 
lamprey (Entosphenus tridentatus) from the Alsea River in Oregon,
and Steelhead trout from Sashin Creek in southeastern Alaska.

The testes were ground in a glender with dry ice prior to extrac-
tion of DNA. The method described by Marmur, modified by addition
of 0.04% deoxycholate to the initial suspension, was employed (see
Appendix I for stepwise procedure) (22, 26). DNA preparations were
stored at 4°C over chloroform or in buffer F. 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 buffer A 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. DNA was prepared for reassociation by the follow-
ing steps: chloroform extraction, centrifugation at 5-10,000 rpm for
about ten minutes, passage through a Metricel GA-6, 45 μ pore size,
cellulose-acetate filter (2) and dialysis. After centrifugation, a thick
emulsion appeared at the interface of the two liquids. The chloroform
phase and the emulsion were separated from the aqueous phase which
contained the DNA by decanting and filtration. The nature of the
emulsion was not investigated.
Labeled DNA was obtained from primary embryonic tissue cultures of steelhead from the Deschutes River. Cells grown in vitro by the method described by Fryer et al. (12) were incubated with C\textsuperscript{14} labeled thymidine (New England Nuclear Corp., 0.22 mC/mg) at 1 μC/ml. The DNA was extracted from pelleted cells by the MUP procedure (6). The DNA was then readsorbed on HAP (hydroxyapatite) (Calbiochem, Bio-Gel HTP), thoroughly rinsed with buffer D and reeluted. No loss in C\textsuperscript{14} counts accompanied this step. The DNA extracted had a specific activity of approximately 4.8 x 10\textsuperscript{5} DPM per OD (optical density) unit. The ratios of absorbance at 260 nm and 230 nm and at 260 nm and 280 nm were lower than those generally observed for the DNA extracted by Marmur's procedure. The labeled DNA was sheared in a Kontes sonicator, catalog number K881440.

Sheared DNA Size

Electron photomicrographs of sheared DNA from steelhead obtained from Sashin Creek 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 M ammonium acetate and shadow cast with carbon or aluminum. The magnification factors of both the electron microscope and the photograph processing were employed in determining the DNA strand length from 8 x 10 in. prints. Single strand lengths of 380 to 540
nucleotides were observed. Because the interpretation of the photographs was difficult, this procedure for estimation of fragment size is not recommended.

The size of the sheared DNA fragments was also estimated from the sedimentation coefficient obtained from band sedimentation velocity experiments (35). The determinations were made in alkaline cesium chloride (3.0 M CsCl + 0.1 M NaOH) 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). Alkaline single stranded sedimentation coefficients were then converted to strand length in nucleotides (37). Values from 380 to 450 nucleotides were obtained which corresponded to the lengths obtained from the electron photomicrographs. The lengths are small enough to avoid gel formations and probably yield uniform kinetics (2).

The sonicated labeled DNA had a single strand length of about 600-1,000 nucleotides (31).

**Melt Curves and DNA Purity Criterion**

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, 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) defined as the temperature at which the absorbance of the sample, 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. The rate of temperature increase was somewhat rapid, producing a lag between the cuvette temperature and recorded temperature; the results, however, were consistent. After connection was made for thermal expansion of the buffer (21), the guanine plus cytosine content was calculated (23); results for steelhead DNA were consistent with values published for rainbow trout, about 43% (11) or a Tm equal to 86.9°C in buffer A.

Melts were made at each stage of one 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). When melts were not run, a ratio of absorbance at 260 nm to absorbance at 280 nm greater than 1.85 and a ratio of absorbance at 260 nm to absorbance at 230 nm greater than 2.3 were adopted as minimum standards for purity (13, 22).
Reassociation of DNA and Hydroxyapatite (HAP) Fractionation

The DNA solutions were placed in cuvettes and treated as above. The cuvette holder accommodated four cells; one position was left 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°C for two minutes—usually as high as 110°C—the holder was removed, unwrapped, and quickly transferred to the sample chamber of the spectrophotometer, which was maintained at the temperature at which the reassociation was to be made, usually 60°C; recording was begun immediately. Approximately 15 minutes elapsed until the temperature in the sample chamber equilibrated. This time had no significant effect on measurements made after extended time. 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 (27). No correction was made for the small absorbance decrease which accompanies the temperature drop initiating the reassociation.

HAP chromatography was employed for fractionating DNA according to the amount of base sequence repetition. Calbiochem Bio
Gel 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. The HAP was boiled twice in buffer D and the finer grains decanted prior to use. A Haake FJ constant temperature circulator provided temperature control. Absorbance readings were made at 320 nm to monitor HAP effluent (3). Fractionation was done as described by Britten and Kohne (4). Slowly reassociating DNA samples were lyophilyzed after fractionation to increase the DNA concentration for subsequent runs.

The second order nature of DNA reassociation was tested by incubating different concentrations of steelhead DNA in buffer F. The amounts of reassociation that had occurred at the same $C_\text{ot}$ value in different runs were compared. $C_\text{ot}$ is the product of the total DNA concentration at time zero and the time elapsed, expressed as (moles of nucleotides)(seconds)(liters)$^{-1}$. The $C_\text{ot}$ 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 $C_\text{ot}$'s and that observed at low concentrations (1 OD) at large $C_\text{ot}$'s, the values observed were quite similar over $C_\text{ot}$ values ranging from 0.1 to 200. Deviations observed with high and low concentrations probably result from the time required for initial temperature equilibration, from
evaporation, from depurination, and from chain scission.

**Thermal Stability Elutions**

DNA whose thermal stability profile was to be examined was adsorbed to HAP in the thermally controlled column described above. The DNA (usually in buffer F, see Appendix III) was made to 0.12 M phosphate with buffer E and passed through the column which was held at 60°C. The perchlorate did not affect the adsorption properties. The column was washed extensively with buffer D at 60°C to remove the single (nonadsorbing) strands of DNA. Execution of the stepwise melt procedure entailed sequential elution of the column with buffer D at progressively higher temperatures. The eluates at each temperature were measured for absorbance, volume, and radioactivity.

**Radioactive DNA Determination**

DNA was precipitated from solution by making the solution ten percent in cold trichloroacetic acid (TCA) and cooled for 15 min in a refrigerator or on ice. The DNA was collected on Schleicher and Schuell B-6, 25 mm, Bac-T-Flex membrane filters; the filters were washed extensively with ten percent TCA. The filters were dried and C\textsuperscript{14} counted in a Packard 2002 Tricarb Scintillation Spectrometer using a scintillation fluid comprised of 5.0 grams of 2, 5-diphenyl-oxazole (PPO) and 0.30 gram of 1, 4-bis-2(4-methyl-5-phenyloxazole)-
benzene (POPOP). Both chemicals were from Packard Instrument Company. An efficiency curve established by counting a series of quenched C\textsuperscript{14} standards (Packard Instrument Company) was used to estimate the activity in disintegrations per minute (DPM) of the sample.

**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, $\frac{A(0)-A(t)}{A(0)}$, against $\log C_{ot}$, where $A(0)$ is the maximum (initial) absorbance, $A(t)$ the absorbance at time $t$, and $A(\infty)$ the absorbance at infinite time (total reassociation). $A(t)$ is used here rather than $C(t)$—concentration—because absorbance ($A$) rather than concentration ($C$) 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 $C_{ot}$ equals 2.752 OD-hours; Britten and Kohne used 1 $C_{ot}$ equal to 2 OD-hours (5).

Native two stranded DNA has a lower absorbance than completely reassociated DNA both because of strand scission and depurination caused by high temperature exposure and because of incomplete reassociation, a result of random reassociation of pairs not sheared at
the same site which produce double stranded products with single stranded "gaps" and loose "ends".

An estimate of absorbance at infinite time may be obtained by two methods. 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 $\frac{A(0)-A(t)}{A(t)-A(\infty)}$ against time, a property of second order kinetics (34) (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 for *E. coli* B, harvested in the stationary phase, were assumed to equal those of other DNAs studied. Values obtained from *E. coli* DNA in different buffers were similar (see Table 1).

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>A</th>
<th>C</th>
<th>B</th>
<th>F</th>
<th>G</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&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.2172</td>
</tr>
</tbody>
</table>

<sup>a</sup>Influenced by evaporation.
A second method of plotting reassociation data was used to illustrate the relative size of repetitious components within the genome. Since the rate at which DNA reassociates at a given \( C_{\text{ot}} \) is proportional to the concentration of sequences reassociating, the slope derived from the reassociation curve [(percent reassociation versus \( \log C_{\text{ot}} \)] plotted against (\( \log C_{\text{ot}} \)], will indicate the relative amount of DNA reassociating at each \( C_{\text{ot}} \). This plot may be more conveniently read as

\((-\log_{10} \text{ repetition frequency}) \) versus (the relative quantity of DNA),

where \((-\log_{10} \text{ repetition frequency}) = \) zero under the peak of unique DNA.
Optical Comparisons of Reassociation and Hybridization Properties

The gross characteristics of ten species of fish are shown by the reassociation \( (C_{ot}) \) curves of the DNAs of these fish (Figures 1A and 1B). Because of the redundancy of some sequences in the genomes of fish, the sheared fragments of DNA in the reassociation mixture are present in varying degrees. Duplex formation of each type of fragment follows second order kinetics. Because of the differences in abundances of the types of fragments, the \( C_{ot} \) curves of the fish reflect the sum of all the different components, their complexities, and relative contributions to the genome. From these \( C_{ot} \) curves, one may estimate the contribution of components repeated to different degrees.

The genomes of all the fish studied consist of two major components, a slowly reassociating fraction representing those sequences present in one or a few copies and a fast reassociating fraction comprised of sequences repeated more than 500 times. The size of the fast component varies considerably among several of the species (Figure 1B; it is constant and relatively larger in the species shown in Figure 1B, salmonids, grayling, and Pacific lamprey. Examination of the \( C_{ot} \) curves shows that only the latter species have a noticeable number of sequences repeated at intermediate (on the order of 100) frequencies.
Figure 1. The reassociation of the DNAs of ten species of fish. The reassociations were done in Buffer F at 60°C using DNA sheared twice at 40,000 psi. The runs were made in 1 mm cuvettes. (A). △ starry flounder at $C_0$ 13.86 OD, ○ ling cod at $C_0$ 9.24 OD, □ carp at $C_0$ 14.17 OD, ▲ buffalo sculpin at $C_0$ 9.38 OD, ● shad at $C_0$ 9.44 OD, and ■ chinook at $C_0$ 9.18 OD. (B). ● whitefish at $C_0$ 10.04 OD, ○ grayling at $C_0$ 16.67 OD, △ lamprey at $C_0$ 15.91 OD, ▲ Deschutes steelhead at $C_0$ 14.07 OD, and □ chinook at $C_0$ 9.18 OD. The chinook is shown in both graphs to provide a reference.
The time required for lowering the temperature of the DNA solution to that of the reassociation from the melting temperature, and the high concentrations of DNA employed prevent accurate descriptions of the heterogeneity and true extent of redundancy at small $C_{ot}$s (on the order of 0.5 $C_{ot}$). In addition, no correction was made for the small absorbancy decrease resulting from secondary structure changes accompanying the initial temperature drop. These restrictions do not affect estimates of the size of the fast and slow components.

The relative sizes of the components are more readily estimated from a plot of the relative reassociation rate on log $C_{ot}$, the derivative of the $C_{ot}$ curve. This curve relates the size of components to the extent of repetition. Using these curves, estimates of the minimum complexity of the species studied were made from the half $C_{ot}$s of the slow fractions. The half $C_{ot}$s of these components were multiplied by the fraction of the total DNA they represented to provide a realistic value for the initial concentration of these components. Examples of these curves are shown in Figures 2, 3, and 4 and a summary of the sizes of fast and slow components and the minimum complexity of these species is given in Table 2.

Some comparisons of the base sequence similarity of these species were made by comparing the $C_{ot}$ curve of a mixture of equal parts of the DNAs from two species to the $C_{ot}$ curves for those species. The three curves were made simultaneously and at quite similar DNA
concentrations. A mixture of DNA from two species, very similar in base sequence, would produce a \( C_{ot} \) curve identical to those of the species being compared. This similarity is illustrated in Figure 2, the comparison of DNAs of Deschutes rainbow and Deschutes steelhead trout.

A theoretical curve expected from a mixture of DNAs from two species which have no sequence similarities was derived from the \( C_{ot} \) curves of the unmixed DNAs of the species being compared. The curve was plotted as one half the sum of the relative absorbancies of the two unmixed DNAs at the same \( C_{ot} \) against twice that \( C_{ot} \). (It takes twice as long for a two-fold diluted sample to reassociate to the same relative extent as does the undiluted sample.) Coincidence of the theoretical and mixture curves is observed at later \( C_{ots} \)s in the comparison of starry flounder with steelhead (Figure 3). The failure of the two curves to coincide at earlier \( C_{ots} \)s probably results from the time required for the initial temperature equilibration rather than from differences in the redundant sequences.

Partial similarity of base sequences is illustrated in Figure 4, the comparison of buffalo sculpin and ling cod DNAs. In this comparison the \( C_{ot} \) curve of the mixture followed a route intermediate to those of the theoretical maximum and the unmixed references. Estimates of the amount of similarity between the species studied were made by comparing the corrected half \( C_{ots} \)s of the mixtures to their
Figure 2. (A) The reassociations of DNAs of Δ Deschutes rainbow at C₀ 13.37 OD, □ Deschutes steelhead at C₀ 13.12 OD, and ○ a mixture of equal parts of both at C₀ 13.50 OD. ● represents a theoretical curve for the mixture of rainbow and steelhead DNA assuming that no sequence similarities exist. The reassociations were run at 60°C in buffer F with DNA twice sheared at 40,000 psi. The runs were made in 1 mm cuvettes. (B) The slopes of the reassociation curves shown in (A) plotted against log C₀t. The assignment of symbols is the same as in (A).
Figure 3. (A) The reassociations of DNAs of □ Deschutes steelhead at $C_0$ 14.07 OD, ○ starry flounder at $C_0$ 13.86 OD, and △ a mixture of equal parts of both at $C_0$ 14.21 OD. ● represents a theoretical curve for the mixture of steelhead and flounder DNA assuming that no sequence similarities exist. The reassociations were run at 60°C in buffer F with DNA twice sheared at 40,000 psi. The runs were made in 1 mm cuvettes. (B) The slopes of the reassociation curves shown in (A) plotted against log $C_0t$. The assignment of symbols is the same as in (A).
Figure 4. (A) The reassocations of DNAs of □ buffalo sculpin at $C_0$ 9.38 OD, ○ ling cod at $C_0$ 9.60 OD, and Δ a mixture of equal parts of both at $C_0$ 9.48 OD. ● represents a theoretical curve for the mixture of sculpin and ling cod DNA assuming that no sequence similarities exist. The reassocations were run at 60°C in buffer F with DNA twice sheared at 40,000 psi. The runs were made in 1 mm cuvettes. (B) The slopes of the reassociation curves shown in (A) plotted against log $C_0$.t. The assignment of symbols is the same as in (A).
Table 2. The size of the fast component and the minimum complexity of the DNAs of ten species of fish and the extent of similarity between them. The half C<sub>0</sub>ts of the non-repeated DNA sequences were corrected to adjust for the contribution to C<sub>0</sub> by the repeated sequences. The number of experiments comprising each average is bracketed. Two experiments are reported for the ling cod-buffalo sculpin comparison and three for the rainbow-starry flounder. These data were obtained from optical comparisons of reassociation and hybridization properties. The reassociations were run at 60°C in buffer F with DNA twice sheared at 40,000 psi. All runs were made in 1 mm cuvettes. The data were obtained from plots such as Figures 2, 3, and 4. DNA concentrations of different runs varied from 9.04 OD to 24.70 OD. An average of five slow chinook-slow steelhead runs at 70°C in buffer F comprise the value reported for the steelhead-chinook comparison. The range of these values was 39.2-80.6%.
<table>
<thead>
<tr>
<th>Species of Fish</th>
<th>Percent Similarity</th>
<th>1/2 $C_{ob}$ of Nonrepeated DNA</th>
<th>Complexity Base Pairs x $10^{-6}$</th>
<th>Percent of DNA in Fast Component</th>
<th>Corrected 1/2 $C_{ob}$ of Nonrepeated DNA</th>
<th>Range of Corrected 1/2 $C_{ob}$</th>
<th>Percent of DNA in Fast Component</th>
<th>Corrected 1/2 $C_{ob}$ of Nonrepeated DNA</th>
<th>Range of Corrected 1/2 $C_{ob}$</th>
<th>Percent of DNA in Fast Component</th>
<th>Corrected 1/2 $C_{ob}$ of Nonrepeated DNA</th>
<th>Range of Corrected 1/2 $C_{ob}$</th>
<th>Percent of DNA in Fast Component</th>
<th>Corrected 1/2 $C_{ob}$ of Nonrepeated DNA</th>
<th>Range of Corrected 1/2 $C_{ob}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starry flounder</td>
<td>34.4</td>
<td>60 [5]</td>
<td>55-60</td>
<td>0.75</td>
<td>8</td>
<td>9</td>
<td>0,0,8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platichthys stellatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ling cod</td>
<td>35.5</td>
<td>65 [3]</td>
<td>60-70</td>
<td>0.81</td>
<td>67,52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ophiodon elongatus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buffalo sculpin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enophrys bison</td>
<td>41.7</td>
<td>60 [4]</td>
<td>60-65</td>
<td>0.75</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprinus carpio</td>
<td>42.8</td>
<td>85 [2]</td>
<td>75-95</td>
<td>1.06</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alosa sapidissima</td>
<td>48.4</td>
<td>80 [3]</td>
<td>70-85</td>
<td>1.00</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grayling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymallus arcticus</td>
<td>60.5</td>
<td>85 [4]</td>
<td>80-95</td>
<td>1.06</td>
<td>84, 32, 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whitefish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prosopium williamsoni</td>
<td>62.0</td>
<td>90 [3]</td>
<td>80-95</td>
<td>1.12</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oncorhynchus tshawytscha</td>
<td>59.9</td>
<td>100 [3]</td>
<td>95-100</td>
<td>1.25</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainbow-steehead</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salmo gairdneri</td>
<td>60.2</td>
<td>130 [5]</td>
<td>130-140</td>
<td>1.62</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lamprey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entosphenus tridentatus</td>
<td>59.8</td>
<td>210 [3]</td>
<td>180-250</td>
<td>2.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
respective theoretical dissimilarity curves. The amount of base sequence similarity between two species was calculated as the difference from unity of twice the ratio of the half Cot of the mixture to the theoretical dissimilarity half Cot. (A ratio of 1.0 corresponds to no similarity, a ratio of 0.75 to 50 percent similarity, and one of 0.5 to complete similarity of base sequences.) Because very little or no similarity existed between species whose complexities differed greatly, no difficulties were encountered for these calculations, and it was possible to cross check these values by comparing the complexity of the mixture to the average of the two species being compared. Table 2 shows the results of the comparisons.

Optical Comparisons of Closely Related Species

Comparisons of the DNAs of closely related species, chinook with steelhead and Deschutes steelhead with Deschutes rainbow, were made to further test the discriminating ability of the technique described in the previous section. Representative $C_0t$ curves (Figures 5 and 8) show that the genome of the chinook is quite similar to that of the steelhead; the Deschutes steelhead and Deschutes rainbow curves coincide. The similarity of these curves makes resolution difficult, but there is a decrease in the rate of reassociation of the chinook-steelhead mixture at later Cots.
Figure 5. 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 buffer F in 1 mm cuvettes with DNA sheared as described in the text.

Figure 6. 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 buffer F in 1 cm cuvettes. The DNA was sheared prior to HAP fractionation.

Figure 7. 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 buffer F in 1 mm cuvettes. The DNA was sheared prior to HAP fractionation.
Figure 8. DNA reassociation curves of Deschutes rainbow $\triangle$ at 13.37 OD, Deschutes steelhead $\square$ at 13.12 OD, and an equal parts mixture of the two $\bigcirc$ at 13.50 OD. The reassociations were run at 60°C in buffer F in 1 mm cuvettes with DNA sheared as described in the text.

Figure 9. DNA reassociation curves of HAP fractionated fast reassociating DNA from Deschutes rainbow $\triangle$ at 1.32 OD, Deschutes steelhead $\square$ at 1.307 OD, and an equal parts mixture of the two $\bigcirc$ at 1.320 OD. The reassociations were run at 60°C in buffer F in 1 cm cuvettes with DNA sheared prior to HAP fractionation.

Figure 10. DNA reassociation curves of HAP fractionated slowly reassociating DNA from Deschutes steelhead $\square$ at 11.38 OD, Deschutes rainbow $\triangle$ at 11.86 OD, and a mixture of equal amounts of the two $\bigcirc$ at 11.50 OD. The reassociation was run at 60°C in buffer F in 1 mm cuvettes with DNA sheared prior to HAP fractionation.
Attempts were made to amplify the differences observed between chinook and steelhead and to resolve differences between the Deschutes steelhead and rainbow. The DNA of all these species was fractionated on HAP to obtain very fast and slowly reassociating components. Samples were lyophylized when more concentrated solutions were required. The reassociating characteristics of the very fast fractions were compared (Figures 6 and 9), but extensive reassociation of the samples and mixtures occurred before the initial temperature equilibration; hence, no differences were resolved. The half Cot of these reactions, a maximum estimate, was of the order of 0.01 Cot, a value which corresponds to a complexity of approximately $10^5$ of nucleotide pairs.

Comparisons of the slowly reassociating DNAs of chinook and steelhead (Figure 7) again displayed the sequence differences observed in the unfractionated comparisons. The slope of the straight line portion of the curve is somewhat less than one might expect from single copy DNA sequences, suggesting the presence of sequences with very low repetition frequencies accompanying the single copy sequences (see Appendix II). Estimates of the extent of similarity from the comparison of the slow components of these species are in Table 2. Although the range of estimates was large, approximately 68% of the non-repeated sequences of chinook were in the steelhead. There were no resolvable differences between the non-repeated portions of the Deschutes rainbow and steelhead DNAs.
Melting Profiles of DNA Hybrids

The thermal stability of DNA hybrids was studied to ascertain the relative extent to which the base sequences of some salmonids had diverged. Small amounts (less than 0.04 OD unit) of C$^{14}$ labeled Deschutes steelhead DNA was added to 7-14 OD of another salmonid species and the mixture heated and incubated. The duplexes formed were adsorbed to HAP and eluted as described earlier in the Materials and Methods section. Because the fast reassociating component produced a duplex whose thermal stability was lower than that of the duplexes from the slow component, the unlabeled DNA was twice fractionated on HAP before using it for the experiment. The resulting slow fraction was about 25% of the total unfractionated DNA. Large concentrations of unfractionated, unlabeled single stranded DNA also formed networks and a precipitate after several days of incubation.

The relatively small concentration of labeled DNA does not contribute significantly to the total DNA concentration, while the extent of reassociation of labeled pieces was minimal and confined to sequences repeated in excess of $10^3$ copies. (The excess of unlabeled DNA was about $10^3$ including the enrichment resulting from fractionation.) It follows that the thermal stability of the labeled DNA describes the degree of complementarity in the hybrid duplexes while that of the unlabeled DNA provides a reference of reassociated strands.
The curves used for these comparisons were adjusted so that they covered a unit area and, therefore, show the relative amounts of DNA (labeled and unlabeled) eluted in each temperature interval, divided by the temperature change of that interval, and plotted against the average temperature of that interval. \[
\frac{(\sum\frac{\Delta \text{DPM}}{\Delta \text{OD}})/(T_2-T_1)}{\text{vs } (T_1 + T_2)/2},\text{ where } T_1\text{ is the temperature at which the previous elution was made, } T_2 \text{ the temperature at which the elution being calculated was made, } \Delta \text{OD or } \Delta \text{DPM the amount of absorbance or label eluted at } T_2, \text{ and } \Sigma \text{DPM or } \Sigma \text{OD the total amounts of label or absorbance at all temperatures greater than the incubation temperature.}\]

Comparison of the unlabeled and labeled duplexes was made by subtracting the temperature at which one half of the total label was eluted from the temperature at which one half of the total absorbance was eluted. This temperature differential provides an index to the amount of mispairing in the hybrid and, therefore, the degree of difference between the two species.

A comparison of slow Deschutes steelhead DNA to labeled unfractionated Deschutes steelhead DNA (Figure 11) served as a reference for other comparisons. A small labeled component displayed lower stability than the majority of the labeled DNA. This component probably resulted from the reassociation of highly repeated sequences of the labeled DNA. This lower stability component tended to amplify
differences in the comparisons of more diverse species such as those shown in Figures 14 and 15.

An attempt was made to discriminate Deschutes rainbow from Deschutes steelhead (Figure 12), but the 1.0° difference obtained from this comparison was less than the 2.25° difference observed in the reference (Figure 11). The precision expected from this type of comparison is indicated by the temperature differentials observed for these two comparisons; this and the other tests described above indicated no differences between the Deschutes rainbow and steelhead. The narrow range of mid-melt temperatures (80.6° to 82°) and the similarity of the shapes of the curves of the unlabeled DNAs substantiate the repeatability and validity of this type of comparisons.

Other profiles to compare of Deschutes steelhead to Sashin Creek steelhead (Figure 13), golden trout (Figure 14), and dolly varden (Figure 15) were made. While a temperature differential of 3° observed in the Sashin Creek steelhead comparison is not substantially different from the Deschutes rainbow or steelhead comparisons, the golden trout at 4.1° and dolly varden at 6.5° are clearly not in this range. Subjectively, the configurations of the curves are quite convincing. Very little of the hybrid of the dolly varden comparison has the thermal stability of the reassociated dolly varden DNA. The golden trout comparison shows two types of DNA hybrid, duplexes whose stability is comparable to that of the golden trout and less stable duplexes.
Figure 11. Melting profiles of duplexes produced by extensively incubating a mixture of slowly reassociating, single stranded Deschutes steelhead DNA (7/OD ml) and single stranded, C$^{14}$ labeled Deschutes steelhead DNA (0.03 OD/ml) at 60°C in buffer F made 0.12 M in phosphate with buffer E. The dotted line follows the melt of the labeled hybrid duplex, the solid line that of reassociated duplexes of slow steelhead DNA. The mid-melt point of the unlabeled DNA was 80.25°C and the temperature differential was 2.25°C.

Figure 12. Melting profiles of duplexes produced by extensively incubating a mixture of slowly reassociating, single stranded Deschutes rainbow DNA (7 OD/ml) and single stranded, C$^{14}$ labeled Deschutes steelhead DNA (0.035 OD/ml) at 60°C in buffer F made to 0.12 phosphate with buffer E. The dotted line follows the melt of the labeled hybrid duplex, the solid line that of reassociated duplexes of slow rainbow DNA. The mid-melt point of the unlabeled DNA was 82.0°C and the temperature differential was 1.0°C.
Figure 13. Melting profiles of duplexes produced by extensively incubating a mixture of slowly reassociating, single stranded Sashin Creek steelhead DNA (14 OD/ml) and single stranded, C14 labeled Deschutes steelhead DNA (0.03 OD/ml) at 60°C in buffer F made to 0.12 M phosphate with buffer E. The dotted line follows the melt of the labeled hybrid duplex, the solid line that of reassociated duplexes of Sashin steelhead. The mid-melt point of the unlabeled DNA was 81.0°C and the temperature differential was 3.0°C.

Figure 14. Melting profiles of duplexes produced by extensively incubating a mixture of slowly reassociating, single stranded golden trout DNA (10 OD/ml) and single stranded, C14 labeled Deschutes steelhead DNA (0.035 OD/ml) at 60°C in buffer F made 0.12 M in phosphate with buffer E. The dotted line follows the melt of the labeled hybrid duplex, the solid line that of reassociated duplexes of golden trout. The mid-melt point of the unlabeled DNA was 80.6°C and the temperature differential was 4.1°C.

Figure 15. Melting profiles of duplexes produced by extensively incubating a mixture of slowly reassociating, single stranded dolly varden DNA (7.5 OD/ml) and single stranded, C14 labeled Deschutes steelhead DNA (0.04 OD/ml) at 60°C in buffer F made 0.12 M in phosphate with buffer E. The dotted line follows the melt of the labeled hybrid duplex, the solid line that of the reassociated duplexes of dolly varden. The mid-melt point of the unlabeled DNA was 80.75°C and the temperature differential was 6.5°C.
IV. DISCUSSION

The information obtained from optical comparisons of the reassociation properties of the DNAs from the ten species studied indicate that fish considered more advanced taxonomically have less information and redundancy in their genomes (Table 2). In starry flounder, ling cod, and buffalo sculpin approximately one-third of the genome was comprised of repeated sequences, while 60 percent was repeated in the salmonids, grayling, and Pacific lamprey; using _E. coli_ B as a reference (8) (see Appendix III), the complexities ranged from $0.75 \times 10^9$ base pairs in the starry flounder to $2.26 \times 10^9$ in the lamprey. Reassociation rate depends somewhat upon the G+C content, but no correction was made for the G+C differences between fish and _E. coli_ (37). It has also been reported that heavy metal ions present during incubation of single stranded DNA may increase the rate of reassociation, and that passing DNA over SE Sephadex prior to incubation removes these ions (32). The rates observed for fish DNA may have been influenced by the presence of these metal ions.

The reasons for variation in the size of the repeated portion are not clear. That the increase in repeated DNA results from additional structural DNA required by organisms with higher chromosome arm numbers is precluded by information from the carp. The carp has about 104 chromosome arms (14) as do many salmonids (30), but a
repeated fraction more similar in size to the repeated portions of the more advanced fish whose arm number is about 50 (38). No effort was made to better describe the repeated portions, but the number of different families of sequences comprising this component and the similarity of the families might provide some insight.

The reduction of information and decrease in the repeated portion parallels a decrease in cellular DNA content accompanying specialization of fish within the same phyletic grouping as was observed by Hinegardner and Rosen (14). Specialization is generally accompanied by loss of structural components and a decrease in plasticity, features consistent with a loss of information (9).

The hybridization information provided by the optical comparison showed extreme base sequence divergence among species considered distantly related taxonomically. While some dissimilarity was expected, the complete absence of similarity observed in some of the comparisons was surprising in that a great number of physical similarities are shared by these organisms. Of particular interest are the comparisons involving salmonids and grayling. The grayling and whitefish are often placed in different families from the steelhead. That the grayling shares more sequences with the whitefish than does the steelhead implies that the common ancestor of the grayling and whitefish existed more recently than did that of the steelhead and whitefish.

Classification systems which place whitefish in the family Salmonidae
should also place grayling there. Accurate information encompassing more species might allow one to "map" their evolution in a manner similar to genetic mapping as was done by Kohne et al. using thermal profiles of hybrids from primates (18).

Attempts were made to increase the resolution of this technique by fractionating the DNA and comparing the slowly reassociating fractions. These attempts did provide consistent resolution between closely related species (chinook and steelhead), but the reliability of the quantitative information did not warrant the effort involved.

In short, these optical comparisons demonstrated that closely related species (buffalo sculpin and ling cod; salmonids and grayling) share a large number of base sequences, but more diverse species share few.

While involving more effort than the optical comparisons, the melting profile of DNA hybrids provide much better resolution and repeatability. The large difference observed from an inter-genus comparison, dolly varden and Deschutes steelhead, was expected, but that between the golden trout and the steelhead was not. Both the Deschutes steelhead and the golden trout possess karyotypes different from that reported for rainbow and steelhead but similar to that of the red banded trout, a trout indigenous to southeastern Oregon. The prevalence of this karyotype led Wilmot to postulate the existence of a widespread complex of fish related intimately to the golden trout (38).
That the Deschutes steelhead and rainbow are closely related to the golden trout was not supported by this comparison. The inability to distinguish between Deschutes steelhead and rainbow indicates that these two fish, which have been treated as separate populations in management programs, may indeed belong to the same breeding population. Alternatively, these fish may belong to separate populations that cannot be resolved by this technique. An isozyme study of these fish by Mark Chilcote (9) did not resolve significant differences between steelhead and rainbows from the main stem of the Deschutes River, the group of fish from which samples for this study were taken.

The different degrees of fractionation between different DNA preparations, the use of varying concentrations of DNA in the incubation mixtures, or variation between experiments, may explain the larger temperature differential observed in the reference (Figure 11) than in the rainbow-steelhead comparison (Figure 12).

Thorgard (32) made similar thermal profiles for salmonids, but used unfractionated, unlabeled DNA and slowly reassociating labeled DNA in his incubations. The conditions he used were 50% formamide + 0.75 M NaCl + 0.075 M Na Citrate, pH 7 at 37°C. These reassociation conditions proved more rigorous because no peak was observed that would correspond to the lower stability peak that appeared when unfractionated DNA was used in experiments reported here. Although his results involved different species of salmonids and are not directly
comparable, they support this approach to the study of base sequence differences among fish.

The most restricting features of this technique are the difficulty involved in preparing DNA with high specific activity and the number of single stranded (unpaired) ends present on the reassociated duplexes that adsorb to the HAP at the start of the melt. An improved technique for radiolabeling DNA in vitro was reported by Commerford (10), involving the iodination of DNA to form 5-iodocytosine using radioiodine. Utilization of this reaction produces labeled DNA which is easier to prepare and which has a much higher specific activity than that made in vivo. Reciprocal tests may be made using DNA fractionated on HAP prior to labeling; the greater specific activity permits a greater ratio of unlabeled DNA to labeled DNA in the incubation mixture. The use of S1 nuclease, a single stranded nucleic acid specific nuclease, prior to the adsorption of duplexes to HAP for the melt, should eliminate any loose ends and badly paired regions. Presumably it would not attack duplexes paired as closely as those of the hybrids being measured. These procedures suggest that increased resolution and reliability may be possible.
V. PERSPECTIVES

The use of reassociation curves in the estimation of similarities between fishes presented several problems. The first involves the measurement of the precise complexity of the species. Unique DNA and DNA possessing some sequences repeated a few times produce quite similar curves. Determination of the curve that represents single copy DNA depends not only on the degree of fractionation and the complexity of the genome but also on reassociation conditions. Some conditions require more homology than do others. The difficulty in estimating the increase in information created by the mixing of DNAs from different sources presents another problem. A 100% increase in information produces a noticeable difference from the reference curve, but obtaining quantitative estimates from a partial increase in information is difficult. Further complications result from the long reassociation at elevated temperatures which produces chain scission and depurination. Although the hybridization data obtained from these optical studies did not prove satisfactory for taxonomic studies, the descriptive information concerning the structural aspects of the genomes of the fish studied is relevant to evolutionary studies.

Of the techniques studied, comparison of the melting profiles of DNA hybrids of two species proved the most applicable to taxonomic
studies. Differences between closely related species were quantitatively measured by methods which could easily be used to catalog species. Although this technique lends itself more readily to examining close relationships, the use of sufficient excesses and varying criteria could provide a means for detecting similarities existing between more distantly related organisms.

The examination and comparison of the base sequences does provide a suitable approach for the systematic study of fish. Quantitative differences were found in the base sequences of closely related species; and, although differences between populations and races were not resolved, the data suggest that, with sufficient refinement of the methodology, such separations might be made.
REFERENCES CITED


APPENDIX I

PROCEDURE FOR DNA EXTRACTION [22, 26]

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 powdered; store frozen.

3. Add (very roughly) 10 g of powdered testis to 125 ml of 0.15 M NaCl + 0.1 M EDTA + 0.04% 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:isoamylalcohol 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 0.1 XSSC. (SSC is 0.15 M NaCl + 0.015 M Na citrate, pH 7).

10. Add 10 XSSC 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, 0.2% in 0.15 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 0.5% 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 0.1 xSCC, add, while stirring constantly, 15 ml of 3.0 M Na acetate + 0.001 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:

\[ C(t) = (k t + \frac{1}{C(0)})^{-1} \text{ or } \frac{C(0)-C(t)}{C(t)} = C(0)k t \]  

(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 reached, the differential equation is:

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

(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.

MIXED COMPONENT SECOND ORDER CURVES

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

\[
\begin{align*}
\frac{dC_1}{dt} & = -k_1 C_1^2, \\
\frac{dC_2}{dt} & = -k_2 C_2^2
\end{align*}
\]

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 16.
Figure 16. Theoretical second order curves for different rate constants \( k \), 1.0, 0.1, 0.01, and 0.001; and two theoretical two-component curves, one of whose rate constants are 0.01 and 0.1 (the two components are present in equal amounts) and the other whose rate constants are 0.01 and 1.
APPENDIX III

BUFFER COMPARISONS

Several buffers were examined to determine the relative rates of reassociation they promoted. *E. coli* B DNA was used because it had no appreciable repeatedness, a convenient complexity, and was easily obtained.

The buffers investigated were buffer A (0.15 M NaCl + 0.015 M Na citrate, pH 7.0), buffer B (0.30 M NaCl + 0.030 M Na citrate, pH 7.0), buffer D (0.12 M PO₄, equimolar in mono- and di-basic Na-PO₄), buffer F (1.0 M NaClO₄ + 0.01 M tris-HCl, pH 7.1), and buffer G (0.5 M NaClO₄ + 0.01 M tris-HCl, pH 7.1). Typical curves of *E. coli* B DNA sheared in buffer A and reassociated in these buffers are shown in Figure 17.

The reassociations of *E. coli* DNA were plotted as percent reassociation on log A₀t. The curves were characterized by the A₀t 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
Figure 17. Reassociations of *E. coli* DNA sheared as described in the text in the buffers,

- buffer A — — •
- buffer B — •
- buffer D — — •
- buffer G — • —
- buffer F — — —
sheared in buffer A, and between DNA slow melted and DNA quick melted prior to reassociation (see Table 3).

Intercept values of DNA reassociation curves indicate that the buffer in which the DNA is sheared significantly effects the reassociation rate (compare Table 3A with 3D). 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 3C and 3D. 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 3A and 3B 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 obvious differences. That of DNA reassociated with buffer B was slightly higher than those of the others; temperature increase caused a decrease in the slopes.

Buffer F increases the reassociation rate about 13 times that of buffer D which is used in HAP chromatography and nearly nine times that of buffer A. The large rate increase produced by the use of buffer F makes it quite valuable for work with eucaryotes whose large genome size necessitates long reassociation. Similarly,
Table 3. Properties of the straight line portion of the plot of percent reassociation on log Cot 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 buffer A and reassociated at 60°C. The buffers are described in the text.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>[cation]</th>
<th>Buffer</th>
<th>A</th>
<th>D</th>
<th>G</th>
<th>F</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>Intercept</td>
<td></td>
<td>0.169(5)</td>
<td>0.249(3)</td>
<td>0.0283(4)</td>
<td>0.0195(3)</td>
<td>0.0299(3)</td>
</tr>
<tr>
<td></td>
<td>Intercept of buffer D (60°C) divided by</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60°C Sheared in SSC</td>
<td>Intercept</td>
<td></td>
<td>1.47</td>
<td>1.00</td>
<td>8.82</td>
<td>12.8</td>
<td>8.34</td>
</tr>
<tr>
<td></td>
<td>Cot 1/2</td>
<td></td>
<td>2.4</td>
<td>3.3</td>
<td>0.53</td>
<td>0.36</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>OD range</td>
<td></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>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td></td>
<td>10.8</td>
<td>10.5</td>
<td>11.5</td>
<td>11.4</td>
<td>12.4</td>
</tr>
<tr>
<td>B.</td>
<td>Intercept</td>
<td></td>
<td>0.242(2)</td>
<td>0.350(1)</td>
<td>0.0430(1)</td>
<td>0.0456(1)</td>
<td>0.0597(1)</td>
</tr>
<tr>
<td></td>
<td>Intercept of buffer D (60°C) divided by</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70°C Sheared in SSC</td>
<td>Intercept</td>
<td></td>
<td>1.03</td>
<td>0.713</td>
<td>5.80</td>
<td>5.47</td>
<td>4.18</td>
</tr>
<tr>
<td></td>
<td>Cot 1/2</td>
<td></td>
<td>3.0</td>
<td>4.0</td>
<td>0.56</td>
<td>0.45</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>OD range</td>
<td></td>
<td>1.8</td>
<td>1.8</td>
<td>1.5</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td></td>
<td>10.1</td>
<td>9.84</td>
<td>9.75</td>
<td>9.05</td>
<td>10.0</td>
</tr>
<tr>
<td>C.</td>
<td>Intercept</td>
<td></td>
<td>0.203(3)</td>
<td>0.246(2)</td>
<td>0.0411(2)</td>
<td>0.0369(1)</td>
<td>0.0509(1)</td>
</tr>
<tr>
<td></td>
<td>Intercept of buffer D (60°C) divided by</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60°C Sheared and reassociated in the same buffer</td>
<td>Intercept</td>
<td></td>
<td>1.23</td>
<td>0.861</td>
<td>6.06</td>
<td>6.75</td>
<td>4.90</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td></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></td>
<td>1.1-1.3</td>
<td>1.3</td>
<td>0.8-1.3</td>
<td>1.2</td>
<td>1.3</td>
</tr>
</tbody>
</table>
Table 3. (Continued)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>[cation]</th>
<th>Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.176(1)</td>
<td>0.163(1)</td>
</tr>
<tr>
<td>Intercept of buffer D (60°C) divided by Intercept</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60°C</td>
<td>Sheared and reassociated in the same buffer</td>
<td>1.42</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>10.7</td>
</tr>
</tbody>
</table>
buffer D 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 buffer A were buffer F at either 60°C or 70°C.