

## КОНСТРУКЦИЯ ВИДА С ТАКСОНОМИЧЕСКОЙ ТОЧКИ ЗРЕНИЯ

Е. С. Смирнов

Из Зоологического института МГУ

### 1. ВВЕДЕНИЕ

Мы считаем полезным предпослать нашему специальному исследованию несколько замечаний общего характера. Дело в том, что систематика в настоящее время занимает особое положение среди биологических дисциплин, и это положение далеко не соответствует удельному весу данной отрасли биологии.

Приходится с сожалением признать, что большинство биологов не имеет сколько-нибудь определенного понятия о задачах систематики как самостоятельной дисциплины. Более того, у многих возникает вопрос, может ли вообще эта наука претендовать на самостоятельное существование.

Между тем всякий биолог, независимо от его специальности, принужден пользоваться категориями системы, — уже по одному тому, что при всяком исследовании, будь оно морфологическое, физиологическое, генетическое, необходимо знать, к какому виду или роду относится объект.

Но разве эти понятия вида, рода, семейства и другие категории системы сами не заслуживают анализа? А система вообще, если иметь в виду ее методологические предпосылки?

Сами систематики мало и неохотно занимаются этими важнейшими вопросами. Мы почти ничего не знаем о теоретическом смысле так называемых высших систематических категорий, начиная от рода и выше. Более посчастливилось элементарной категории — виду, которому посвящена обширная литература. Однако большинство исследований, посвященных виду, характеризуется крайней нестрогостью. В этой области царит методологическая и фактическая путаница. Действительно, до настоящего времени отсутствует мало-мальски строгое определение вида. Говоря о виде, обычно указывают на его наследственную стойкость и на высокую степень сходства между различными поколениями. Но ведь сходство — понятие относительное: оно теряет смысл, если не сопровождается объективным критерием. Старое определение, которое утверждает, что сходство особей одного вида того же порядка, как сходство детей с родителями, конечно, не выдерживает критики. Что же касается стойкости видовых признаков, то ведь и родовые признаки отличаются наследственной стойкостью. Кроме того, такие группы, как биотип или чистая линия, тоже являются, по мнению генетиков, наследственно стойкими единицами.

Существуют два наиболее распространенных мнения о задачах систематики, и оба они таковы, что прямо или косвенно отрицают самостоятельную ценность этой дисциплины.

Многие считают, что исследования систематического характера имеют только вспомогательное, подсобное значение. С этой точки зрения задача систематика состоит в том, чтобы составить удобный

для обозрения каталог форм растений и животных. Этот каталог представляет собой как бы справочник, который позволяет нам узнать, к какому виду, роду, семейству и т. д. относится интересующий нас в том или ином отношении объект. Ясно, что данная точка зрения исключает какую бы то ни было ценность систематического исследования, кроме узко практической.

Другие авторы, желая произвести систематику в ранг научной дисциплины, стремятся превратить ее в филогенетику. Речь идет о том, чтобы представить существующую систему в форме генеалогического дерева. Такая филогенетическая система мыслится как гигантская генеалогическая схема, на которой располагаются систематические группы в порядке их кровного родства.

Критика этой концепции давалась различными авторами, в том числе и мной (6)<sup>1</sup>. Мы не имеем возможности рассмотреть здесь этот вопрос, но должны заметить, что сведение системы к генеалогии тоже равносильно отрицанию самостоятельного значения систематики. Попробуем теперь охарактеризовать главные систематические проблемы.

Основное положение, которое, вероятно, без труда примет большинство биологов, имеющих знакомство с той или иной отраслью систематики, заключается в признании реальности систематических групп. Это означает, что названные группы действительно существуют в природе и могут быть вполне объективно охарактеризованы нашими систематическими категориями вида, рода, семейства и т. д. Задача систематики заключается в том, чтобы правильно констатировать эти группы и изучить закономерность их строения. Констатация и характеристика групп приводят нас к построению естественной системы. Существующую иерархию соподчиненных систематических категорий мы и считаем естественной системой.

Проблема строения систематических категорий включает вопросы закономерности распределения и соотношения признаков, теорию их таксономического значения, вопрос о взаимоотношении компонентов систематических групп и др. Не входя в дальнейшие подробности, мы отсылаем читателя к другим нашим статьям. Кроме того, мы рассчитываем в дальнейших главах настоящей работы показать на конкретных примерах существо систематических проблем.

Мыслима и действительно существует диаметрально противоположная точка зрения на систему. Мы имеем в виду тот взгляд, что никаких систематических групп в природе нет, и реальны только индивиды. Систематические категории с данной точки зрения представляют собой произвольные сочетания индивидов, причем принцип этого сочетания определяется не объективными явлениями природы, но вкусом исследователя или какими-либо специальными задачами, им поставленными. Если занять такую позицию, то вопрос о закономерностях строения систематических групп, конечно, потеряет свой смысл.

Не вступая здесь в дискуссию со сторонниками данного взгляда, мы считаем, что установление закономерного строения систематических групп—лучший способ доказательства правильности нашей точки зрения.

## 2. ЭЛЕМЕНТАРНЫЙ ВИД И ПРОБЛЕМА СХОДСТВЕННЫХ ОТНОШЕНИЙ

В своих предыдущих работах, посвященных теории системы (6—11), мы имели дело с анализом высших систематических категорий, начиная с рода и выше. Далее мы пришли к убеждению, что ряд ос-

<sup>1</sup> Из других более новых работ см. Любищев (†).

новых систематических проблем значительно легче поддается разрешению, если в качестве материала взять элементарную группу.

Таким элементом считается вид. Однако это справедливо лишь при условии, что вид далее не дробится на естественные группы. При современном уровне наших знаний надо считаться с тем, что в пределах вида очень часто заключается большее или меньшее количество подчиненных единиц (подвидов, или рас), которые далее дробятся только на индивиды.

Существует, однако, достаточно большое число видов, не дробящихся на подчиненные единицы. Вид, послуживший объектом настоящего исследования, относится как раз к этой категории; поэтому в дальнейшем, говоря о строении вида, мы будем иметь в виду элементарный вид, не распадающийся на расы. Наши выводы относятся к таким видам или же в равной степени к расам сложных видов.

Как было сказано выше, мы исходим из того, что вид есть некоторая реальность, объективно существующая в природе. В противном случае было бы нелепо говорить о его строении.

Какие свойства вида подлежат изучению с точки зрения систематики? В этом заключается кардинальный вопрос. Генетик изучает наследственные свойства индивидов, принадлежащих данному виду, и в результате синтеза полученных данных может судить о генетическом составе вида. Эколог рассматривает отношение особей к окружающей обстановке и в результате приходит к экологической характеристике вида. Морфолога интересуют вопросы архитектоники—строения тела, физиолога—жизненные отправления и т. д. Нередко исследование производится на одном или немногих представителях данного вида. Полученные результаты в этом случае обычно распространяются на весь видовой коллектив в силу того, что изученные экземпляры достаточно представляют данный вид. ●

Всеми перечисленными исследованиями, к которым надо еще прибавить данные других биологических дисциплин, систематическая проблема вида отнюдь не решается. Остается очень важный вопрос о сходственных отношениях особей, слагающих вид. Устанавливая закономерности, определяющие степень и характер сходства особей вида, мы переходим в область основной систематической проблемы. Эту задачу можно формулировать несколько иначе, говоря о законах изменчивости, существующих в пределах вида. Мы все же предпочитаем первую формулировку, так как понятию изменчивости придается в настоящее время различный смысл, что вносит неопределенность и создает путаницу.

Понятие сходства более объективно, но, конечно, и оно требует анализа. Ему и посвящена значительная часть нашей работы: мы стремимся к тому, чтобы сделать это понятие вполне объективным.

Степень сходства устанавливается посредством сравнения интересующих нас объектов. В пределах коллектива; каковым является вид, подлежащий нашему изучению, сравнение каждого индивида с каждым натолкнуло бы нас на слишком большие трудности, так как потребовало бы огромного числа манипуляций. Для сравнительного изучения «коллективного предмета» (термин Гурвича) существует более рациональный и простой способ: мерилom можно взять некий эталон, или единицу измерения, общую для всех компонентов коллектива.

Вообще говоря, мы свободны в выборе мерилa. Например, можно взять эталоном любую из особей данного вида и по сравнению с ней судить об особенностях прочих. Однако этот выбор произволен—у нас нет основания для того, чтобы избрать именно эту особь, а не какую-нибудь другую.

Между тем имеется вполне объективное мерилло, которое обладает очень важными преимуществами. Мы имеем в виду среднюю арифметическую. Как известно, средняя обладает тем свойством, что сумма квадратических отклонений всех вариантов какого-либо признака от нее является минимальной. Это означает, что любой индивид нашей группы, взятый в качестве эталона, дальше отстоит в отношении данного признака от совокупности прочих индивидов, чем средняя арифметическая. Если бы мы взяли для характеристики группы по этому признаку вместо средней какую-нибудь особь, то совершили бы большую ошибку. Разумеется, и средняя не дает точной характеристики группы, но она ближе всего к истине. Итак, средняя арифметическая является вполне объективным мериллом при определении сходства; она лучше характеризует группу, чем любой из индивидов, и, кроме того, значительно облегчает технику вычисления при оценке сходства. Все это вместе взятое вполне оправдывает ее применение для нашей цели.

Разберем теперь вопрос о свойствах индивида, в отношении которых мы производим сравнение. В этой операции мы принуждены опираться на отдельные признаки, так как суммарная оценка будет слишком неточна. Но признаки разнообразны. Для совершенно точного и объективного учета степени сходства мы должны были бы взять их все. Проще всего взять признаки строения. Но почему мы стали бы пренебрегать признаками физиологического порядка? Нет никакого основания думать, что физиологическая индивидуальность выражена слабее морфологической. То же следует сказать о признаках экологического характера, генетического и т. д.

Для каждого из признаков, независимо от его характера, мы можем найти среднюю арифметическую и по ней оценивать индивидуальные особенности. Правда, не все признаки могут быть выражены количественно, но для оценки качественных признаков существует свой критерий, который с математической стороны адекватен средней. Благодаря этому единство сравнения не нарушается.

Таким образом, мы получим совокупность средних арифметических всех признаков нашего вида. Эта совокупность обозначается как видовой тип. Видовой тип мы и примем в качестве эталона, или нормы сравнения. Это—некоторый идеальный средний индивид. Он характеризует наш коллектив с большей степенью точности, чем любая фактически существующая особь, так как совокупность средних, очевидно, будет обладать основным свойством средней одного признака, о чем мы говорили выше.

Далее, пользуясь типом, мы без труда можем охарактеризовать любой индивид нашей группы. Для этого нужно найти разности между размером каждого признака ( $a$ ) и соответственной видовой средней ( $M_a$ ). Эти разности берутся по абсолютной величине, не считаясь со знаком: нам важно знать отклонение от типа, а не направление его. Затем было бы естественно взять сумму величин  $|a - M_a|$  и разделить ее на число изученных признаков ( $n$ ). Выражение  $\sum |a - M_a| : n$  дает понятие о том, насколько данный индивид в среднем отклоняется от видового типа.

Однако, в это рассуждение необходимо внести существенную поправку. Дело в том, что нельзя сравнивать различные признаки непосредственно. Допустим, что мы сравниваем отклонения от типа в двух признаках: длине тела и высоте головы. Ввиду гораздо большего размера первого признака, сравнительно незначительное отклонение длины тела от нормы может резко превысить самое крайнее отклонение во втором признаке.

Беря такое соотношение без поправки, мы резко повысим удельный вес первого признака по сравнению со вторым, что недопусти-

мо по логическим соображениям. Необходимость поправки станет еще очевиднее, если мы будем сравнивать два признака разного наименования, например, размер и вес тела. Разумеется, нельзя складывать сантиметры с граммами и, чтобы выйти из этого затруднения, надо признаки сделать сравнимыми. Это производится при помощи среднего квадратического отклонения ( $\sigma_a$ ). Разделив выражение  $|a - M_a|$  на величину  $\sigma_a$  того же признака, мы получим отвлеченное число, которое называется нормированным отклонением. Оно показывает, во сколько раз отклонение особи в отношении признака  $a$  превышает среднее квадратическое отклонение, характерное для данного признака. Такие выражения делают все признаки вполне сравнимыми, независимо от их категории и наименования.

Таким образом, выражение  $\frac{|a - M_a|}{\sigma_a} : n$  показывает величину отклонения некоторой особи от видового типа, выраженную отвлеченным числом. Пользуясь этой формулой, мы, казалось бы, можем установить, как распределяются особи, слагающие вид, по степени типичности. Однако на этом пути мы встречаем новое и более серьезное затруднение, которое будет разобрано в следующей главе.

### 3. ПРИНЦИП ИСЧЕРПЫВАЮЩЕЙ ХАРАКТЕРИСТИКИ

Для того чтобы получить исчерпывающе точное представление о степени сходства двух индивидов, нам нужно учесть все их признаки. Дело, конечно, не меняется и в том случае, когда мы характеризуем особей через посредство видового типа, определяя степень типичности каждой из них.

Сколько же признаков имеет индивид?

Небольшое размышление показывает, что число их бесконечно велико. Более того, даже в пределах одного органа или его части мы можем констатировать бесконечное множество признаков. Взяв в качестве примера бедренную кость какого-нибудь млекопитающего, мы можем провести любое количество сечений, перпендикулярных ее оси. Сечений будет тем больше, чем меньше расстояния между ними. В пределах каждого сечения мы точно так же можем провести сколько угодно большое число диаметров, постепенно уменьшая углы между ними. Аналогичным образом мы сталкиваемся с бесконечностью и в пределах небольшого сектора этого сечения, поскольку нельзя поставить никаких границ уменьшаемости углов. Все это — признаки.

Но, благодаря этому, задача исчерпывающей характеристики становится неразрешимой: исчерпать все признаки невозможно, следовательно, нельзя говорить и о точном установлении сходства.

Логический выход из этого положения на первый взгляд можно найти, базируясь на исследовании Heinske (2). Если оценить степень сходства между особями вида (у него — расы) по одному признаку, то получится обычная нормальная кривая распределения или какое-либо из ее видоизменений. Это означает, что, наряду с особями, близкими к средней, будут и такие, которые резко отличаются от нее. Прибавим еще один признак и нормируем отклонения особей от средних, а затем для каждой особи определим среднее нормированное отклонение, взятое по первому и второму признакам.

Heinske считает, что полученные таким образом числа  $\left(\frac{a_1}{\sigma_1} + \frac{a_2}{\sigma_2}\right) : 2$  дадут меньший размах изменчивости, чем числа  $a_1 : \sigma_1$  или  $a_2 : \sigma_2$ . Он объясняет это тем, что ни одна особь, принадлежащая к одной и той же расе, не дает резких отклонений от нормы во всех признаках

одновременно. Напротив, резкие отклонения нескольких, немногих, признаков компенсируются тем, что громадное большинство остальных признаков лишь слабо отклоняется от нормы. Совпадение резких отклонений даже в двух признаках у всех особей является по Heinske невероятным. Увеличивая число признаков, мы должны будем, естественно, прелучать все большее и большее сближение особей в отношении их характеристик, т. е. чисел  $\sum \frac{a - M_a}{\sigma_a} n$ , так как совпадение одинаково резких отклонений в нескольких признаках будет еще менее вероятным. Когда число признаков будет бесконечно велико, все особи окажутся одинаково типичными. Иначе говоря, в этом случае  $\sum \frac{a - M_a}{\sigma_a} n = const.$ , при  $n \rightarrow \infty$ .

Этот результат можно выразить иначе: изменчивость нормированных отклонений в  $n$  признаках подчиняется в пределах особи тому же самому закону нормального распределения, что и изменчивость отклонений, взятых по одному признаку, но в пределах всей группы особей вида.

Свои соображения Heinske подтверждает некоторыми примерами, правда, немногочисленными: особи, резко уклоняющиеся в одном признаке, оказываются у него более типичными с возрастанием числа  $n$ . Аномалии отдельных признаков компенсируются большим числом признаков.

Если бы Heinske оказался прав, то выход из упомянутого выше затруднительного положения был бы найден. В самом деле, если бы удалось показать, что все особи элементарного вида равнотипичны, то проблема закономерности их распределения была бы решена раз навсегда. А в этом случае всех особей одного вида пришлось бы считать по существу одинаковыми, так как разница между ними заключалась бы только в том, что одни и те же величины отклонений падали бы у различных особей на различные признаки.

В одной из своих статей я уже высказал свое несогласие с Heinske. Однако, ввиду важности вопроса, мы должны подвергнуть его более строгому анализу, которым сейчас и займемся.

В качестве объекта для решения вопроса о конструкции вида с таксономической точки зрения мы взяли одно из обычных двукрылых насекомых нашей фауны — *Dolichopus plumipes* Scop.—вид, принадлежащий к семейству *Dolichopodidae*. Наша пробная группа была составлена из 100 самок этого вида, собранных в окрестностях Москвы. Мы ограничились особями одного пола по той причине, что данный вид характеризуется значительным половым диморфизмом.

По отношению к признакам, подлежащим изучению, мы применили тот же метод пробной группы, выбрав некоторые особенности строения скелета. Крыло особенно удобно для нашей цели, так как легко поддается характеристике: достаточно разветвленное жилкование дает ряд легко устанавливаемых точек отсчета.

Группа признаков была взята следующим образом. Через две точки мы провели ось крыла (рис. 1). Первая из них — пересечение медиальной жилки с задней поперечной ( $F$ ), вторая — впадение медиальной в край крыла ( $J$ ). На эту ось мы спроектировали ряд других точек: впадение базальной поперечной жилки в субкостальную ( $A$ ), основание вспомогательной ( $B$ ), впадение радио-медиальной в медиальную ( $C$ ), впадение радиальной  $R_1$  в край крыла ( $D$ ), пересечение радио-медиальной с дискоидальной жилкой ( $E$ ), впадение дискоидальной в задний край крыла ( $G$ ), впадение радиальной  $R_{2+3}$  в

передний край крыла ( $H$ ) и аналогичное впадение  $R_{4+5}$  ( $I$ ). Кроме того, мы провели перпендикуляр к оси через вершину крыла ( $K$ ), т. е. точку, наиболее удаленную от основания крыла.

В результате мы получили две группы признаков. Первая—отрезки оси крыла, определяемые проекциями названных выше точек. Мы обозначаем их по направлению от основания крыла к его вершине через  $a, b, c, \dots, j$ . Вторая группа признаков—отрезки перпендикуляров между названными точками и осью крыла:  $k, l, m, \dots, s$ . Сумму отрезков  $a, b, c, \dots, j$  мы называем длиной крыла  $S$ .

К указанным 20 признакам мы присоединили впоследствии еще 21, беря расстояния между различными точками пересечения жилок и впадения их в край крыла. Эти отрезки следующие:  $AD, EG, CD, DF, EF, HI, AB, BC, IJ, FG, DE, BE, FH, GJ, GH, AC, GI, DG, EH, DH, CF$ .

Из всей массы признаков одна лишь длина крыла была взята по абсолютной величине. Остальные измерения относились к размеру  $S$ .

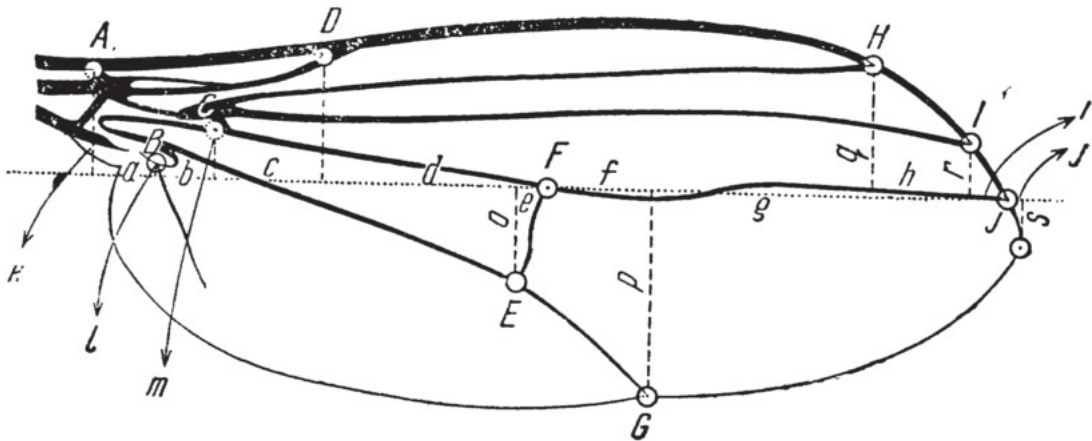


Рис. 1

Таким образом, мы имели дело в дальнейшем с индексами типа  $a:S$ . Причину такого выбора признаков мы разъясним в следующей главе.

Первая задача нашего анализа заключалась в том, чтобы установить, как влияет количество учтенных признаков на индивидуальные характеристики, т. е. числа  $\sum \frac{|a - M_a|}{\sigma_a} : n$ . Для этой цели нам пришлось нормировать все признаки. Каждое индивидуальное отклонение в некотором признаке  $a$  выражалось, таким образом, числом  $\left(\frac{a}{S} - M_{\frac{a}{S}}\right) : \sigma_{\frac{a}{S}}$ . Эти числа брались по абсолютной величине, что уже было мотивировано выше. Среднюю величину таких чисел для  $n$  признаков, характеризующую индивид, мы обозначим  $I_n$ . Эта средняя по двум признакам индивида будет  $I_2$ , по трем— $I_3$  и т. д.

Ради экономии места мы не приводим ни средних арифметических, ни средних квадратических отклонений рассмотренных признаков крыла, ибо нас интересуют прежде всего корреляции признаков.

Согласно Нейске, числа  $I$ , резко разнящиеся индивидуально при учете немногих признаков, должны постепенно сближаться с возрастанием количества признаков  $n$ , стремясь в пределе к тождеству. Но как найти критерий сближения  $I$ ? Самый простой способ заключается в том, чтобы при данном  $n$  построить вариационный ряд из чисел  $I$  и затем найти среднее квадратическое отклонение  $\sigma_{I_n}$ . Чем ближе друг к другу  $I$ , тем меньше должно быть среднее квадратическое отклонение, являющееся, как известно, мерой дисперсии. При наличии тождества чисел  $I$  эта мера будет равна нулю:  $\sigma_{I_\infty} = 0$ .

Пользуясь этим методом, мы последовательно вычислили  $\sigma_I$  для одного, двух и т. д. признаков, кончая 41-м.

Таблица 1. Порядок признаков:  $b, d, f, h, i, j, o, S, a, c, e, g, AD, k, l, n, m, p, q, r, s, G, CD, DF, EF, HI, AB, BC, IJ, FG, DE, BE, FH, GJ, GH, AC, GI, DG, EH, DH, CF$

Число признаков $n$	$\sigma_{I_n}$	Число признаков $n$	$\sigma_{I_n}$	Число признаков $n$	$\sigma_{I_n}$	Число признаков $n$	$\sigma_{I_n}$
1	.6422	12	.2018	22	.1969	32	.1760
2	.4436	13	.2020	23	.1961	33	.1752
3	.4017	14	.1978	24	.1936	34	.1741
4	.3447	15	.2000	25	.1896	35	.1744
5	.2951	16	.2051	26	.1901	36	.1765
6	.2485	17	.2137	27	.1857	37	.1780
7	.2259	18	.2067	28	.1808	38	.1764
8	.2151	19	.2051	29	.1812	39	.1776
9	.2115	20	.2028	30	.1796	40	.1792
10	.2041	21	.2015	31	.1776	41	.1776
11	.1974						

В табл. 1 приводятся величины  $\sigma_I$  при постепенно возрастающем  $n$ . Последовательность, в которой мы брали признаки, указана в заголовке таблицы. Ради наглядности мы приводим еще рис. 2, где показана графически связь между количеством использованных признаков и величиной  $\sigma_I$ . Между обеими переменными существует определенное соотношение. Первоначальное увеличение  $n$  действительно дает резкое уменьшение  $\sigma_I$ . Прибавление второго признака, например, уменьшает величину  $\sigma_I$  почти в  $1^{1,2}$  раза. Пять признаков дают по сравнению с одним сокращение дисперсии более чем вдвое, восемь признаков—в три раза и т. д. Однако прибавление 12-го признака неожиданно дает увеличение  $\sigma_I$ , а 13-й признак—еще некоторое повышение. При  $n=14$  кривая опять понижается, но следующие 3 признака приводят к резкому повышению. В результате 17 признаков дают почти тот же уровень кривой, как и первые восемь: прибавление 9 признаков практически не уменьшило дисперсии.

Дальнейший рост  $n$  все время дает уменьшение  $\sigma_I$ , которое, правда, идет довольно медленно. Это можно констатировать для следующих 17 признаков, т. е. до  $n=34$ . (Впрочем, здесь имеет место небольшое нарушение этой закономерности: прибавление 26-го признака приводит к незначительному возрастанию  $\sigma_I$ , которое далее компенсируется при увеличении  $n$ .) Положение резко меняется, когда  $n$  достигает 35. Начиная с 35-го признака, кривая уже не обнаруживает систематического приближения к оси абсцисс. Ординаты колеблются вокруг некоторого среднего уровня, причем колебания эти сами по себе совершенно незначительны. Величина  $\sigma_I$  при  $n=41$  в точности та же, что при  $n=31$ . Таким образом, прибавление к 31 признаку еще 10 несколько не уменьшило дисперсии.

Наш результат, очевидно, находится в полном противоречии с точкой зрения Нейпске. Во-первых, явно несправедливо то, что прибавление новых признаков к уже имеющимся непременно уменьшает дисперсию,—может иметь место и обратное явление. Можно сказать лишь в общей форме, что при сравнительно небольших пределах роста числа признаков уменьшение дисперсии случается чаще, чем увеличение ее. Во-вторых, кривая изменения  $\sigma_I$  отнюдь



не стремится к пересечению с осью абсцисс, ибо  $\sigma_1$  не стремится к нулю. Она асимптотически приближается к прямой, параллельной оси  $x$ , которую приблизительно можно определить уравнением  $y = 0,1766$  (беря в качестве  $y$  среднее из значений  $\sigma_1$  при  $n$  от 31 до 41, т. е. в тех пределах, где наша кривая уже не приближается к оси  $x$ ).

Отсюда вытекает очень важный вывод. Если мы используем известную категорию признаков (в данном случае признаков, характеризующих пропорции крыла), то нет надобности для исчерпывающей характеристики сходственных отношений индивидов брать

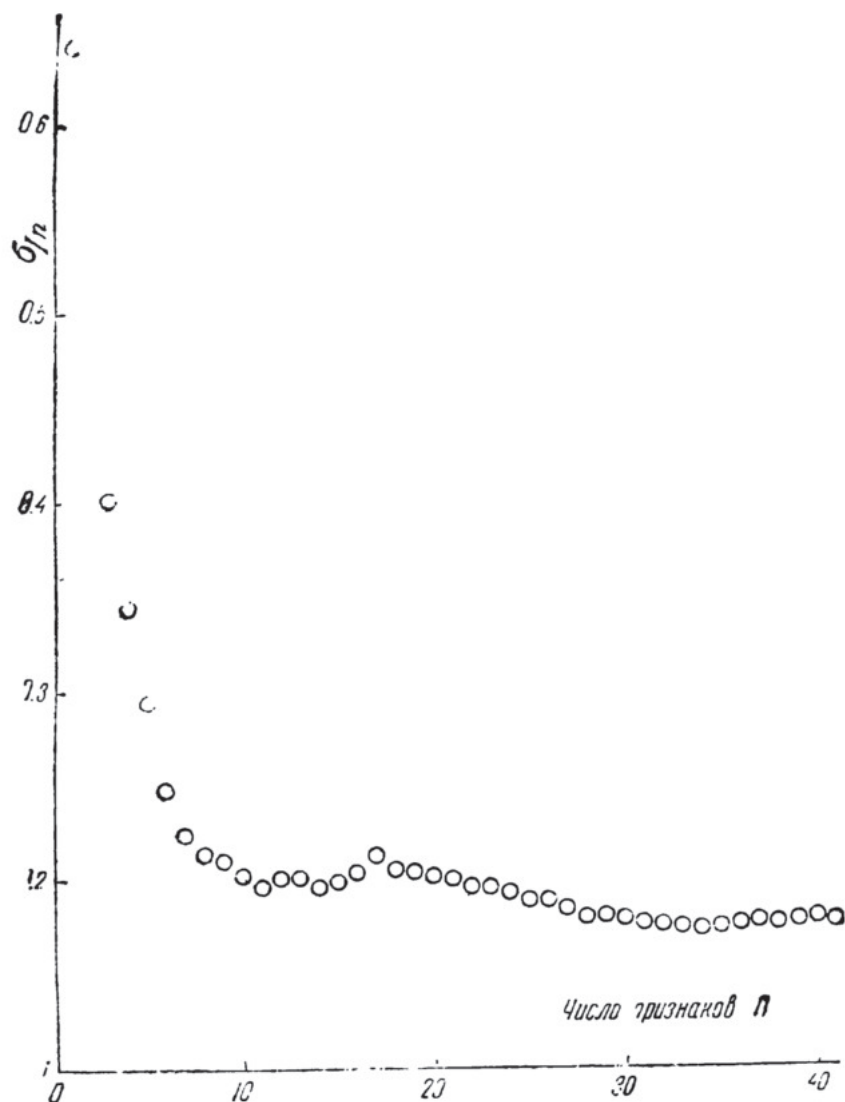


Рис. 2

бесконечно большое количество признаков. Мы можем ограничиться известной пробной группой их, которая не только конечна, но и сравнительно невелика. Мы могли бы характеризовать пропорции крыла не четырьмя десятками признаков, но сотнями, тысячами и вообще любым количеством их, но такое увеличение  $n$  означало бы простую тавтологию, поскольку среднее отклонение от типа становится постоянным уже при трех десятках признаков данной категории.

Наш вывод обнаруживает закономерность, которую мы можем назвать принципом исчерпывающей характеристики. Он решает поставленную нами проблему: каким образом возможно точное определение сходства при бесконечно большом числе признаков? Анализ этого принципа посвящено дальнейшее изложение.

#### 4. СВЯЗАННОСТЬ ПРИЗНАКОВ

Центр тяжести вопроса заключается в том, что признаки неравнозначны при установлении сходственных отношений, а неравнозначность зависит от наличия между ними коррелятивной связи.

Heincke был бы совершенно прав, если бы признаки были вполне независимы друг от друга. В самом деле, при этом условии нормированные отклонения должны комбинироваться в пределах индивида по принципу случайности. Резкое отклонение индивида в каком-либо признаке несколько не влекло бы за собой резких отклонений в других отношениях. Компенсация сильных отклонений слабыми имела бы место, числа  $l$  непрерывно сближались бы друг с другом, и величина  $\sigma_l$  стремилась бы к нулю по мере возрастания числа признаков  $n$ .

Таблица 2. Корреляции крыловых признаков первой группы.  
Все признаки типа  $a: S$ , кроме  $S$

	$a$	$b$	$c$	$d$	$e$	$f$	$g$	$h$	$i$	$j$	$S$
$a$	—	-.417	-.053	.147	-.243	-.108	-.126	.046	-.199	-.479	187
$b$	-.417	—	-.095	.263	.185	-.071	-.092	.027	-.011	.133	.046
$c$	-.053	-.095	—	.412	-.166	-.059	.144	-.175	-.224	.017	.237
$d$	.147	.263	.412	—	-.156	.000	.471	.005	-.004	-.048	.193
$e$	-.243	.185	-.166	-.156	—	-.591	.231	-.005	.104	.031	.198
$f$	-.108	-.071	-.059	.000	.591	—	-.490	-.119	-.039	.006	.101
$g$	-.126	-.092	.144	.471	.231	.490	—	-.311	-.180	-.102	.102
$h$	.046	.027	-.175	-.005	.05	.119	.311	—	-.031	-.027	.097
$i$	-.199	-.011	-.224	-.004	.104	-.039	.180	.031	—	.152	.039
$j$	-.479	.133	.017	.048	.031	.006	-.102	-.027	.152	—	.110
$S$	187	.046	.237	.193	.198	.101	.102	.097	-.039	.110	—

Таблица 3. Корреляции крыловых признаков второй группы.  
Все признаки типа  $k: S$ , кроме  $S$

	$k$	$l$	$m$	$n$	$o$	$p$	$q$	$r$	$s$	$S$
$k$	—	.883	.954	.881	-.092	.376	.166	-.162	.394	-.251
$l$	.883	—	.919	.759	-.168	.325	-.312	-.221	.430	-.250
$m$	.954	.919	—	.881	.108	.376	-.149	-.180	.440	-.288
$n$	.881	.759	.881	—	.025	.330	-.046	-.076	.397	-.231
$o$	-.092	-.168	.108	.025	—	.358	.138	.043	-.125	.071
$p$	.376	.325	.376	.330	.358	—	.103	-.112	.185	-.106
$q$	-.166	-.312	-.149	.046	.138	.103	—	.496	-.201	.316
$r$	-.162	-.221	-.180	-.076	.043	-.112	.496	—	-.187	-.029
$s$	.394	.430	.440	.397	-.125	.185	-.201	-.187	—	-.221
$S$	-.251	-.250	-.288	-.231	.071	-.106	-.316	-.029	-.221	—

Представим себе теперь обратный случай, когда все признаки находятся в функциональной зависимости друг от друга. В этом случае коэффициент корреляции, как известно, измеряется единицей:  $r = \pm 1$ . Но если так, то всякому резкому отклонению в одном признаке должны сопутствовать столь же резкие отклонения и во всех остальных (при наличии прямолинейной связи признаков, которая имеет место в нашем случае).

В первом случае все признаки необходимы. Прибавление любого признака прибавляет нечто к характеристике индивида, делая ее более точной, позволяет с большей точностью определить расстояние ( $l$ ) индивида от типа.

Во втором случае нам совершенно достаточно взять лишь один признак, притом любой из всей массы их. В силу полной корреляции остальные  $(n - 1)$  признаки дадут тот же самый результат. Индивид, резко уклоняющийся в одном отношении, будет так же резко отклоняться в остальных.

Таким образом, количество взятых признаков не будет влиять на точность характеристики, каждое  $I$  останется постоянным при любом  $n$ , и дисперсия ( $\sigma_I$ ) не будет меняться.

В действительности признаки связаны попарно самой различной степенью корреляции, так что величина  $r$  колеблется в широких пределах, между 0 и  $\pm 1$ . Отсюда логически вытекает, что значение признаков для характеристики индивидов фактически меньше, чем в нашем первом случае, но больше, чем во втором.

В табл. 2 и 3 мы приводим все попарные коэффициенты корреляции первой и второй групп крыловых признаков. Признаки взяты в относительной форме (типа  $a : S$ ). В обеих таблицах даны также корреляции всех признаков с длиной крыла  $S$ . Как видно из этих таблиц, признаки связаны попарно очень различной корреляцией. Наряду с высокой степенью зависимости ( $r = + .954$ ) мы встречаем исчезающе слабую связь ( $r = + .025$ ), когда величина  $r$  меньше средней ошибки.

В табл. 3 обращает на себя внимание группа из 4 признаков —  $k, l, m, n$ , образующая тесно связанный комплекс, внутри которого  $r$  не падает ниже .759.

Т а б л и ц а 4. Порядок признаков:  
 $k, l, m, n$

Число признаков ( $n$ )	$\sigma_{I_n}$
1	.5609
2	.5345
3	.5406
4	.5183

Если мы выделим эту группу и произведем вычисление  $\sigma_I$ , последовательно прибавляя по одному признаку, то получим результат, приведенный в табл. 4. Мы видим здесь, что возрастание числа признаков лишь очень слабо уменьшает величину  $\sigma_I$ . Сравнивая  $\sigma_{I_4}$  с  $\sigma_{I_1}$ , мы находим, что прибавление трех признаков ( $l, m, n$ ) уменьшило  $\sigma_I$  лишь на .0426, что составляет приблизительно 7,6% по отношению к  $\sigma_{I_1}$ . Кроме того, прибавление третьего признака  $m$  не только не уменьшает  $\sigma_I$ , но даже увеличивает ее довольно значительно.

Аналогичная операция для признаков  $b, d, f, h$  (табл. 1) дает уменьшение на 46,3%, т. е. в несколько раз большее.

Отсюда вытекает, что признаки далеко не равноценны, и дело не только в их количестве, но также и в качестве. Признаки, связанные сильно ( $k, l, m, n$ ), дают гораздо меньшее понижение  $\sigma_I$ , чем слабо связанные ( $b, d, f, h$ ). Это нетрудно понять на основании сказанного выше. Сильно связанные признаки в значительной степени повторяют друг друга. Отклонения, особо резкие или очень слабые в одном из признаков такой группы, лишь немного сглаживаются другими. Благодаря этому величина  $\sigma_I$  незначительно реагирует на увеличение  $n$ .

Это и заставило нас взять крыловые признаки не в абсолютной, а в индексной форме. Абсолютные размеры отрезков крыла очень

тесно связаны друг с другом, так как все они находятся в сильнейшей зависимости от величины крыла. Использование их представляло бы мало смысла. Выражая же наши признаки в относительной форме, мы освобождаем их от этой зависимости. Действительно, табл. 2 и 3 показывают, что признаки индексные связаны с длиной крыла слабо.

Уменьшение  $\sigma_I$ , которое мы констатировали столько раз, не трудно понять. Прибавление нового признака, слабо связанного с предыдущими, сглаживает резкие отклонения индивидов (чисел  $I$ ) от типа, так как слабость связи означает, что резкие отклонения  $(n+1)$ -го признака не совпадут с резкими же отклонениями первых  $n$  признаков. На резкие отклонения  $I_n$  падут как раз многочисленные мелкие отклонения  $(n+1)$ -го признака.

Но что означает увеличение  $\sigma_{I_n}$  в случае прибавления  $(n+1)$ -го признака?

Допустим, что признаки  $x, y$  совершенно не связаны друг с другом, и  $r_{xy} = 0$ . В таком случае  $\sigma_{I_2}$  будет гораздо меньше, чем  $\sigma_{I_1}$ . Прибавим теперь третий признак  $z$ , который слабо связан с  $x$ , но сильно с  $y$ . Что произойдет с  $\sigma_I$ ? В силу высокой корреляции  $z$  и  $y$  первый из этих двух признаков будет усиливать значение резких отклонений второго. Резкие отклонения по признаку  $y$ , сглаженные было наличием  $x$ , снова выступят на сцену, так как будут в значительной степени совпадать с сильными же отклонениями  $z$ . Очевидно, что в результате прибавления признака  $z$ ,  $\sigma_{I_3}$  будет больше  $\sigma_{I_2}$ . Эта операция увеличила удельный вес признака  $y$ , в то же время ослабив значение  $x$ .

Теперь представим себе, что мы вычислили  $\sigma_{I_n}$  для большой группы признаков. В качестве примера обратимся к табл. 1 и обратим наше внимание на том моменте изменения  $\sigma_I$ , когда  $n$  достигло 13. Прибавление 14-го признака заметно уменьшает  $\sigma_{I_{13}}$ , но далее, по мере привлечения признаков  $l, m, n$ , наблюдается сильный и постоянный рост величины  $\sigma_I$ .

Как уже было сказано, эти четыре признака очень тесно связаны друг с другом, образуя некоторый комплекс. Беря их всей четверкой, мы искусственно создаем перевес этой группы признаков над остальными.

Итак, мы видим, что возрастание  $\sigma_I$  при увеличении  $n$  определяется двумя факторами. С одной стороны, это — теснота связи, т. е. величина показателя  $r$ . Если бы в предыдущем примере признаки  $k, l, m, n$  были связаны функционально, то перевес их над предыдущими сказался бы еще больше: отклонения индивидов в этих признаках были бы подчеркнуты еще сильнее, и мы получили бы более резкое увеличение  $\sigma_I$ .

С другой стороны, для последнего необязательно наличие очень сильной связи. Допустим, что мы вычислили  $\sigma_I$  для десяти признаков, совершенно не связанных. Если теперь прибавить еще несколько признаков, не зависящих от первой группы ( $r = 0$ ), а друг с другом связанных слабо (например,  $r = 0,2$ ), то мы все же, в конце концов, должны получить увеличение  $\sigma_I$  по сравнению с  $\sigma_{I_{10}}$ . В самом деле, эти новые признаки в своих отклонениях хотя и слабо, но все же повторяют друг друга. Но это означает, что уже при  $n = 12$  объединенные 11-й и 12-й признаки получили некоторый перевес над первыми десятью. Конечно, он был бы большим при наличии более высокой корреляции.

Таким образом, помимо абсолютной величины коэффициентов корреляции, для изменения  $\sigma_{I_n}$  немаловажное значение имеют также и их относительные размеры, т. е. отклонения  $r$   $(n+1)$ -го,  $(n+2)$ -го

и следующих признаков от коэффициентов корреляции, связывающих попарно уже использованные  $n$  признаков. В силу этого приходится ставить вопрос о дисперсии коэффициентов корреляции между  $(n+1)$ -м признаком и первыми  $n$ .

Какой функцией определяется зависимость  $\sigma_I$  от обоих факторов — величины коэффициентов  $r$  и их дисперсии? Этот вопрос требует математической обработки. Не будучи математиком, я не беру на себя его решение.

## 5. ТАКСОНОМИЧЕСКОЕ ЗНАЧЕНИЕ ПРИЗНАКОВ

Итак, критерий  $\sigma_I$  обладает еще одним важным свойством: он возрастает в том случае, когда мы, прибавляя новые признаки к ранее взятым, искусственно создаем перевес некоторой группы признаков над другими. Значение этого свойства  $\sigma_I$  будет особенно ясным из следующего примера.

Допустим, что к 13 первым признакам нашей табл. 1 мы прибавляем не только  $k, l, m, n$ , но и ряд других признаков, тесно связанных с этой четверкой. Можно, например, провести ряд прямых, параллельных этим перпендикулярам, и взять затем их отрезки, заключенные между точками переднего края крыла и их проекциями на ось (рис. 3,  $ab, a_1b_1, a_2b_2, \dots$ ).

Беря в качестве новых признаков достаточно большое количество таких отрезков, мы, в конце концов, почти сведем к нулю роль первых 13 признаков. Действительно, поскольку между  $k, l, m, n$  существует очень тесная связь, совершенно очевидно, что и промежуточные отрезки будут не менее тесно с ними связаны; поэтому наш критерий  $\sigma_I$  будет отражать почти исключительно особенности части крыла, заключенной между  $k, n$ , передним краем крыла и соответственным отрезком крыловой оси. Но это будет означать пренебрежение всеми остальными компонентами крыла. Между тем наша цель — дать исчерпывающую характеристику особей по этому органу в целом. Неправомерность такой операции совершенно ясна. И вот как раз увеличение  $\sigma_I$  в таком случае предохраняет нас от ошибки, автоматически ее регистрируя.

Все сказанное приводит нас к очень важному вопросу о выборе признаков и их таксономическом значении. Таксономическая ценность признака, очевидно, зависит от того, насколько он улавливает характерные свойства индивида, отличающие его наилучшим образом от других индивидов той же группы. Наиболее ценное в определении сходства — установление важнейших различий!

Как мы видели выше, все признаки равноценны в том случае, когда они совершенно независимы друг от друга, так как здесь каждый прибавляет одинаково много к характеристике особи. Равноценность остается и в том случае, когда связь признаков абсолютная, или функциональная. Однако в этом случае, в противоположность первому, необходимым является лишь один (любой) из признаков, так как он одновременно представляет и все остальные. Прибавление даже второго признака не имеет смысла, так как означает тавтологию — все элементы независимости уже исчерпаны первым признаком.

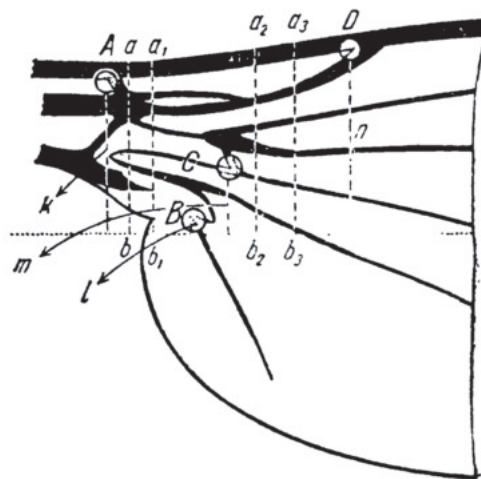


Рис. 3

Рациональная характеристика заключается в том, чтобы учесть все те элементы независимости, которые имеют место в действительности, т. е. в тех промежуточных случаях, когда признаки связаны разнообразным образом и неодинаково отражают свойства индивида. Таксономическая ценность признаков будет тем выше, чем лучше выполняют они эту задачу. Показателем же будет служить большая или меньшая степень уменьшения нашего критерия  $\sigma_I$ .

Для того чтобы облегчить решение интересующего нас вопроса, мы рассмотрим сначала некоторые фиктивные случаи.

1. Допустим, что орган распадается на ряд участков или зон. Между ними никакой связи нет, но внутри каждого участка все части связаны функционально. Это означает, что корреляция интерзональных признаков равна нулю, а интразональных — единице.

Очевидно, что для таксономической характеристики органа нам нужно учесть все зоны. Если мы возьмем по одному признаку от каждой зоны, то получим характеристику, которая будет исчерпывающей и в то же время не будет содержать тавтологии.

Если теперь от одной из наших зон взять второй признак, то пропорциональность представительства нарушится, так как эта зона искусственно получит перевес над прочими. В этом случае  $\sigma_I$  обнаружит приращение. Ошибка увеличится, если мы от данной зоны возьмем не два, а большее количество представляющих ее признаков.

Будет ли ошибкой, если мы от каждой зоны возьмем по два или больше, но во всяком случае по равному количеству представляющих зоны признаков? Нет, пропорциональность не нарушится, но такая операция будет означать тавтологию. Характеристика останется достаточной и исчерпывающей, но нельзя сказать, чтобы двойное, тройное и т. д. представительство было необходимым, ибо оно просто излишне. Критерий  $\sigma_I$  остается без изменения, а это и показывает наличие тавтологии.

2. Задача усложнится, если мы примем, что внутри каждой зоны связь не функциональная, а коррелятивная, причем теснота этой связи варьирует в широких пределах (межзональная связь попрежнему отсутствует).

В этом случае нам нужно исчерпать элементы независимости внутри каждой зоны. Следовательно, интразональные признаки должны быть избраны так, чтобы остальные, не вошедшие в число избранных, давали бы увеличение  $\sigma_I$  или оставляли ее без изменения. Таким образом, каждая зона будет охарактеризована критерием  $\sigma_{I_k}$ , где  $k$  — число избранных признаков, которое будет варьировать от одной зоны к другой.

После этого нам придется учесть все зональные особенности, т. е. найти  $\sigma_I$  для признаков, избранных по всем зонам вместе. Все эти признаки необходимы, так как между признаками разных зон по предположению корреляции нет. В результате мы получим исчерпывающую характеристику органа  $\sigma_{I_{k_1 + k_2 + k_3 + \dots + k_p}}$ , где  $p$  — число независимых друг от друга зон, а буквы  $k_1, k_2, \dots, k_p$  обозначают числа признаков в отдельных зонах.

3. Допустим теперь, что интразональная связь равна 1, а между зонами имеется корреляция, причем  $r$  варьируют от 0 до  $\pm 1$ .

Очевидно, что здесь мы, во-первых, должны взять по одному (любому) признаку, представляющему свою зону, а далее из полученных  $p$  признаков извлечь все элементы независимости, т. е. отобрать  $(p-x)$  признаков с таким расчетом, чтобы остальные  $x$  не изменяли бы  $\sigma_I$  или давали бы ее возрастание. Характеристика  $I_{p-x}$  и будет исчерпывающей. В то же время она не страдает тавтологичностью.

4. Наиболее общим является тот случай, когда связь интразональная, равно как и интерзональная, варьирует от 0 до  $\pm 1$ , не достигая этих предельных значений. Мы все же сохраняем право говорить о зонах, так как допускаем, что интразонально признаки связаны сильнее, чем интерзонально.

Однако для составления исчерпывающей характеристики эти зоны не имеют значения. Здесь не может быть строгого представительства отдельных зон, так как они не замкнуты внутри себя. Задача исчерпывающей характеристики решается так, что мы избираем  $s$  признаков независимо от зон, и далее, всякий  $(s+1)$ -й признак либо увеличивает  $\sigma_{I_s}$ , либо оставляет эту величину без изменения.

## 6. РАСПРЕДЕЛЕНИЕ КОРРЕЛЯТИВНЫХ СВЯЗЕЙ

Естественно возникает вопрос: что имеет место в действительности? Каковы закономерности связи признаков в пределах органа и между органами? В качестве ответа приведем несколько примеров.

Ильин (3), изучавший корреляции белых пятен у морских свинок, находит, что все точки депигментации (так он называет места образования белых пятен) могут быть распределены в 7 групп. Эти группы содержат точки, связанные сильно ( $r > 0,5$ ), тогда как сами они «почти совершенно не коррелированы между собой. Беря группы более крупные, можно свести их к трем (область головы и шеи, уха, все остальное туловище), причем «эти области, безусловно, независимы одна от другой».

Мы можем согласиться с автором в том, что корреляции пятен распределяются неравномерно и что действительно можно выделить известные зоны более сильной связанности. Но автор неправ, утверждая, что его зоны абсолютно не связаны: корреляцию ниже 0,45 он считает отсутствием связи, с чем никак нельзя согласиться. Картина, полученная им, соответствует нашему общему случаю, когда имеется более сильная связь внутри зон и ослабленная между зонами.

Интересно установленное Ильиным соответствие между топографической близостью признаков и степенью их связанности. Это наблюдение гармонирует с предложенным Перлем «правилом соседства». Однако такое соответствие выражено в материалах Ильина не слишком строго и допускает много исключений.

В одной из старых работ, посвященной крылу *Drosophila funebris* Fbr. (12), мы дали большое число корреляций между абсолютными длинами отрезков жилок крыла. Обзор этих данных привел нас к заключению, что крыло *Drosophila* с точки зрения корреляций лучше всего рассматривать как векториальное поле. В нем можно констатировать известные направления сильной и слабой корреляции. Направление поперек крыла (перпендикулярное его длинной оси) встречает признаки сильно связанные, тогда как направление, параллельное оси, характеризуется слабой связью признаков. Таким образом, части одной и той же продольной жилки связаны слабо, отрезки же разных продольных, но лежащие в одной поперечной зоне, коррелированы сильно.

Уже после напечатания названной статьи мы попробовали найти закон распределения коэффициентов корреляции  $r$  с целью выделения двух возможных групп: интразональных, с одной стороны, и интерзональных — с другой. На основании 170 коэффициентов мы построили кривую распределения; хотя она и оказалась двувёршинной, но вторая вершина выражена слабо и статистически недостоверна. Это приводит нас к выводу, что установленные поперечные зоны сильно

коррелированных признаков все же достаточно связаны и друг с другом.

Терентьев в своей интересной работе (14), посвященной анализу корреляции признаков у *Rana ridibunda* Pall., приходит к заключению о наличии «плеяд» сильно коррелированных признаков<sup>1</sup>. Им было изучено свыше двух десятков признаков (промеров), характеризующих пропорции тела, и получен 231 коэффициент  $r$ . Распределение величин  $r$  оказалось бимодальным, причем одна из вершин (слабая корреляция!) приходится на «межплеядные» коэффициенты, а другая (большие  $r$ ) — на внутрплеядные. Ввиду того, что связь всех признаков с общими размерами тела маскирует истинные связи между частями тела, автор обратился к парциальной корреляции, исключив сначала влияние длины тела. Распределение парциальных  $r$  вновь бимодально, причем один максимум соответствует межплеядным признакам ( $r$  от 0,0 до 0,5), другой — внутрплеядным ( $r$  от 0,6 до 0,9). Между обеими группами существует узкая полоса смешанных признаков ( $r$  от 0,5 до 0,6). Интересно отметить, что внутрплеядные связи относятся только к скелетным признакам (их 10); прочие признаки — «кожные», как их обозначает Терентьев.

Далее, автор анализирует найденную плеяду тоже при помощи парциальной корреляции. Мы вернемся к этому ниже, а сейчас подчеркнем, что межплеядные связи у *Rana ridibunda* достаточно высоки, ибо  $r$  может достигать до 0,5. Наша «зона» в общем не совпадает с «плеядой» Терентьева, но сущность дела от этого не меняется.

Корреляции, вычисленные нами к настоящей работе, тоже дают некоторый материал по интересующему нас вопросу.

Обратимся к нашей табл. 2, где даны попарные корреляции индексных признаков  $a, b, c, \dots, j$ , а также связи с длиной крыла  $S$ . Первое, что бросается в глаза, — это общая слабость связей. Если ограничиться признанием только тех корреляций, когда  $r$  превышает среднюю квадратическую ошибку *minimum* втрое, то лишь  $r \geq 0,276$  следует признать имеющими реальное значение. Таких коэффициентов в нашей таблице имеется только 7 из 45, т. е. 15,5%. Взяв признаки в индексной форме и освободившись от влияния общих размеров крыла, мы действительно добились резкого повышения независимости признаков друг от друга.

Все 7 коэффициентов отрицательны. Это естественно, так как при постоянной сумме всех проекций ( $S$ ) увеличение одних, вероятнее всего, связано с уменьшением других. Влияние соседства на интенсивность связи сказывается в том, что 6 коэффициентов, имеющих реальное значение, относятся к соседним признакам. Вместе с тем оно далеко не абсолютно. Например, признак  $a$  реально связан с соседним  $b$ , но, кроме того, с наиболее отдаленным  $j$ , тогда как промежуточные признаки не обнаруживают заметной связи с ним. Признак  $d$  действительно связан с соседом  $c$ , но также и с  $g$ , в то время как промежуточные  $e, f$  от него не зависят. Мы не говорим уже об очень многочисленных случаях, когда признак не связан ни с одним из соседей.

Помимо влияния соседства, мы не можем констатировать какие-либо иные закономерности в распределении корреляции. Теперь поставим целью выбрать среди наших признаков такую группу, чтобы внутри нее не наблюдалось реальных связей. Иными словами, между любой парой их не должно быть корреляции выше 0,275. Задача эта допускает два решения: одна группа будет иметь состав  $a, c, e, g, i$ , другая —  $b, d, f, h, j$ . К каждой из этих групп мы имеем право при-

<sup>1</sup> Этот термин употреблялся им и ранее (13).



бавить еще  $S$ . В первой группе мы избежали всякого соседства, во второй — допустили лишь признак  $i$ , который, однако, не связан со своими соседями  $h, j$ .

Очевидно, мы не можем в данном случае говорить о каких-либо зонах связанных признаков вдоль крыла.

Иную картину мы наблюдаем в отношении ординат точек крыла. Среди признаков  $k, l, \dots, s$  резко выделяется своей взаимной связанностью четверка  $k, l, m, n$ . К ней примыкают менее тесно связанные с этой четверкой  $p$  и  $s$ , — последний, несмотря на свое отдаленное положение.

Если бы мы захотели выделить из признаков  $k, l, \dots, s$  группу не связанных взаимно, то пришлось бы ограничиться лишь тремя:  $o, q, s$ .

Итак, мы можем констатировать прикорневую зону крыла, которая в поперечном направлении сжимается и расширяется в общем как целое — пропорционально. Однако этого нельзя сказать о продольном направлении:  $a$  связан с  $b$  посредственно, а проекция  $c$  и вовсе не связана с ними.

Может возникнуть еще один вопрос — о связи обеих групп  $a, b, c, \dots$  и  $k, l, m, \dots$  друг с другом. Мы не вычисляли всех возможных коэффициентов корреляции между признаками этих групп, имеющиеся же в нашем распоряжении говорят о большом разнообразии связей. В среднем связи эти невелики. Однако, если бы мы пожелали увеличить комбинацию независимых признаков за счет  $k, l, m, \dots, s$ , то пришлось бы воспользоваться всего одним признаком  $o$ , который можно присоединить к группе  $b, d, f, h, i, j$  (но не к  $a, c, e, g, i$ ), прибавление, отнюдь не пропорциональное числу новых признаков.

В общем распределение корреляций позволяет говорить лишь о частичной зональности крыла, к тому же слабо выраженной. Другими словами, перед нами тот же общий, случай связи признаков, о котором говорилось выше.

Помимо крыловых корреляций, мы анализировали также связи компонентов конечностей *D. plumipes* — передних и средних. В качестве признаков были взяты длины бедра, голени и каждого из 5 члеников лапки. (Ноги брались всегда правые.) Это — та же самая пробная группа особей, в которой мы изучали крыловые признаки. Корреляции абсолютной величины всех компонентов ноги оказались очень высокими. Табл. 5, относящаяся к передним конечностям, дает об этом понятие.

Таблица 5. Корреляции отделов передней конечности. Признаки взяты в абсолютной форме

	$F$	$T$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
Femur ( $F$ )	—	.970	.902	.864	.821	.814	.788
Tibia ( $T$ )	.970	—	.880	.876	.821	.815	.804
Tarsus I ( $t_1$ )	.902	.880	—	.869	.821	.800	.778
» II ( $t_2$ )	.864	.876	.869	—	.924	.843	.859
» III ( $t_3$ )	.821	.821	.821	.924	—	.929	.854
» IV ( $t_4$ )	.814	.815	.800	.843	.929	—	.903
» V ( $t_5$ )	.788	.804	.778	.859	.854	.903	—

Мы видим, что величина  $r$  варьирует в очень узких пределах, от 0,788 до 0,970. Все коэффициенты при нашем числе вариантов, равном 100, вполне надежны. Аналогичная картина наблюдается и для средних ног, где  $r$  варьирует от 0,777 до 0,957.

При такой всеобщей высокой корреляции трудно обнаружить какие-нибудь закономерности в распределении связи. Влияние общих

размеров органа на отдельные его части маскирует более тонкие закономерности.

Ввиду этого мы и здесь прибегли к индексным признакам, отнеся длины компонентов к общей длине ноги  $L$ , т. е. к сумме всех измеренных частей. Результат вычисления корреляций представлен в табл. 6 и 7.

Таблица 6. Корреляции отделов передней конечности. Признаки типа  $F:L$

	$F$	$T$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
Femur ( $F$ )	—	.286	— 236	— .645	— 584	— .522	— 434
Tibia ( $T$ )	.286	—	— .519	— .451	— .513	— .462	— 362
Tarsus I ( $t_1$ )	— 236	— .519	—	— .014	— .058	— .136	— .057
» II ( $t_2$ )	— .645	— 451	— 014	—	.553	427	.280
» III ( $t_3$ )	— .584	— 513	— .058	.553	—	.566	.378
» IV ( $t_4$ )	— 522	— .462	— .136	.427	.566	—	.502
» V ( $t_5$ )	— 434	— .362	— .057	.280	378	.502	—

Таблица 7. Корреляции отделов средней конечности. Признаки типа  $F:L$

	$F$	$T$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
Femur ( $F$ )	—	.021	— .409	— .278	— 381	— .404	— .187
Tibia ( $T$ )	.021	—	— .185	— .515	— .490	— .355	— 423
Tarsus I ( $t_1$ )	— .409	— 185	—	— .117	— .198	— .047	— .083
» II ( $t_2$ )	— .278	— .515	— .117	—	.463	197	.072
» III ( $t_3$ )	— .381	— .490	— 198	463	—	.388	.272
» IV ( $t_4$ )	— .404	— .355	— .047	.197	.388	—	.350
» V ( $t_5$ )	— .187	— .423	— 083	.072	.272	.350	—

Здесь мы легко можем зафиксировать характерные закономерности корреляции, в общем одинаковые для обеих пар ног. Это особенно касается знака  $r$ . Положительная корреляция имеет место, во-первых, между бедром и голенью (впрочем, ничтожная для средних ног и слабая для передних). Затем, ясно выраженную группу положительно связанных признаков образуют членики тарсуса от II до V. Мы отчеркнули на таблицах соответственные области  $r$ , где все коэффициенты положительны. Следовательно, относительное удлинение одного из четырех члеников влечет за собой удлинение остальных трех. Это правило действительно для обеих пар ног.

С названной четверкой признаков отрицательно связаны бедро, голень и I членик лапки: это — правило без единого исключения, действительное как для передних, так и для средних ног. Следовательно, относительное удлинение одного из первых трех отделов ноги связано с укорочением последних четырех. Таким образом, по знаку корреляции мы с уверенностью можем разделить конечность на две зоны, изменение которых происходит в противоположном

направлении. Что касается тесноты связи, интересующей нас более всего, то здесь о сколько-нибудь ясной зональности говорить не приходится: распределение  $r$  по его абсолютной величине дает пеструю картину.

В дополнение сказанного представляет интерес разбор связей между компонентами различных конечностей. Соответственные  $r$  были нами вычислены, причем оказалось, что наиболее тесно (при том положительно) связаны соответственные части ног, особенно же  $t_3 - t_3$ ,  $t_4 - t_4$  и  $t_5 - t_5$ . (Эти связи, однако, не выше 0,6.) Следовательно, можно утверждать, что соответственные отделы разных пар конечностей изменяются от индивида к индивиду в общем согласованно. Кроме того, важно подчеркнуть, что обе пары конечностей (нужно думать, и все три) с точки зрения корреляции представляют одну систему: величины  $r$ , связывающих разные конечности одной стороны, в общем того же порядка, что и в пределах одной конечности.

Резюмируя эту главу, мы должны признать, что в пределах органа или иной части организма нет достаточно резко выраженных и отграниченных друг от друга зон корреляции. Вследствие этого мы не можем строго провести принцип представительства зон при вычислении  $\sigma_I$ : приходится решать задачу исчерпания элементов независимости для органа в целом.

## 7. ПРАКТИКА ВЫБОРА ПРИЗНАКОВ

До сих пор мы обсуждали таксономическое значение признаков главным образом с теоретической стороны. Но спрашивается, как его определить практически и как дать ему численное выражение?

Разберем этот вопрос на нескольких примерах. Первый относится к нашим крыловым (индексным) признакам  $a, b, c, \dots, r, s$ , к которым мы прибавили еще длину крыла  $S$ . В общем мы рассматриваем здесь 20 признаков.

Критерий  $\sigma_I$  при  $n = 20$  равен 0,2080. Для того, чтобы определить таксономическое значение каждого из 20 признаков, мы будем последовательно исключать каждый из них и вычислять  $\sigma_I$  для совокупности остальных 19 признаков. Разности  $\sigma_{I_{20}}$  и получаемых при исключении каждого признака  $\sigma_{I_{19}}$  покажут нам влияние этих признаков на  $\sigma_{I_{20}}$ .

Результат вычислений приведен в первой графе табл. 8.

Здесь мы находим в одном столбце величины  $\sigma_{I_{19}}$ , а в другом — разностей ( $\sigma_{I_{19}} - \sigma_{I_{20}}$ ). Эти разности отнюдь не одинаковы. Прежде всего надо отметить, что 5 из них отрицательны: прибавление одного из признаков  $k, l, m, n, q$  к прочим девятнадцати вызывает не уменьшение  $\sigma_I$ , но его увеличение. Остальные 15 признаков уменьшают  $\sigma_I$ , но далеко не в одинаковой степени. Например,  $r$  дает уменьшение на .0018, а  $o$  — на .0113, т. е. в 6,8 раз больше предыдущего.

Нужно ли исключить из нашей системы все 5 признаков, увеличивающих  $\sigma_I$ ? Из дальнейшего будет видно, что такая операция повредила бы решению нашей задачи — рационального выбора признаков.

Мы исключим сначала признак  $m$ , который дает наибольшее увеличение  $\sigma_I$ . После этого мы снова вычислим влияние каждого из оставшихся 19 признаков на  $\sigma_{I_{19}}$ . Во второй графе табл. 8 приводятся  $\sigma_{I_{18}}$  и ( $\sigma_{I_{18}} - \sigma_{I_{19}}$ ), получаемые при последовательном выключении признаков  $a, b, c, \dots, j, k, l, n, \dots, s, S$  по одному. В этом случае мы, как правило, получаем другие разности, чем в первом, но признаки

$k, l, n, q$  попережнему увеличивают  $\sigma_I$ . Наибольшее увеличение дает  $l$ , который мы и выключаем на этот раз, получая  $\sigma_{I_{18}} = 1956$ .

Новая проверка таксономического значения признаков  $a, b, c, \dots, j, k, n, \dots, s, S$  при исключенном  $l$  дает опять иную картину (графа третья). Особенно важно, что, помимо  $k$  и  $n$ , увеличивавших  $\sigma_I$  и в прошлый раз, появляется один новый признак —  $r$ , изменяющий наш критерий в том же направлении, впрочем, очень слабо. Кроме того, признак  $s$  вообще не влияет на величину  $\sigma_I$ : разность  $(\sigma_{I_{17}} - \sigma_{I_{18}})$  равна нулю.

Исключение следующего признака  $n$  приводит нас к небезынтересному результату (графа четвертая). Мы видим, что признак  $k$  теперь играет уже положительную роль в характеристике, давая уменьшение  $\sigma_I$ . Признак  $q$  продолжает увеличивать  $\sigma_I$ , а  $r$ , дававший увеличение, теперь не влияет на величину критерия:  $\sigma_{I_{16}} - \sigma_{I_{17}} = 0$ . Напротив,  $s$  дает обратную картину. Мы исключаем на этот раз признак  $q$ , как дающий максимальный рост  $\sigma_I$ , после чего  $\sigma_{I_{16}} = 1904$ .

В результате этой операции (графа пятая) признак  $s$  сохраняет свою отрицательную роль в характеристике, а признак  $r$  приобретает исчезающе малое положительное значение; кроме того, обнаруживается отрицательное действие  $g$  и тавтологичность признака  $d$  (ибо  $\sigma_{I_{15}} - \sigma_{I_{16}} = 0$ ).

Теперь нам приходится исключить признак  $g$ , после чего (графа шестая) остается один лишь  $s$ , увеличивающий  $\sigma_{I_{15}}$ . Признак  $d$  из тавтологического стал полезным для характеристики.

Последующее исключение  $s$ , уменьшающее величину  $\sigma_I$  до 1880, дает окончательный результат: ни один из оставшихся 14 признаков не увеличивает  $\sigma_{I_{15}}$ .

Изложенный пример позволяет сделать ряд важных заключений. Самое важное следствие то, что признаки неравноценны для характеристики и что самое понятие таксономической ценности имеет относительное значение.

Взяв наудачу 20 признаков одной категории, мы убедились в том, что значительная часть их не только не увеличивает степени точности характеристики индивидов, но даже, напротив, уменьшает ее, давая неверное понятие о сходстве сравниваемых объектов. В результате исключения этих шести признаков нам удалось уменьшить величину  $\sigma_I$  с 2080 до 1880, т. е. на 0,200, что составляет 9,6%. Отрицательная роль трех из этих признаков ( $l, m, n$ ) отмечалась уже нами выше. Здесь мы можем сделать аналогичное утверждение в более строгой форме. Названные признаки, будучи тесно связанными, теряют в своей таксономической ценности. Они создают искусственный перевес представляемой ими зоны над другими компонентами крыла.

Было бы, однако, неправильным вместе с признаками  $m, l, n$  исключить и четвертый член этой группы —  $k$ . В таком случае данная зона совсем не была бы представлена в характеристике крыла. Мы можем констатировать на примере этого признака относительность таксономической ценности. При наличии  $m, l$  и  $n$  признак  $k$  давал увеличение  $\sigma_I$  а при их отсутствии дает уменьшение. Один и тот же признак, в зависимости от наличной комбинации других, может быть то полезен, то вреден для характеристики.

Об этой относительности говорит вся наша таблица. Признаки  $g$  и  $s$ , казавшиеся полезными вначале, оказались вредными в дальнейшем. Признак  $d$  был бесполезен при комбинации  $a, b, c, e, f, g, h, i, j, k, o, p, r, s, S$ , но приобрел известное значение при окончательном наборе

Таблица 8. Таксономическое значение крыловых признаков

Признаки типа  $a : S$

Признаки	$\sigma_{I_{20}} = \cdot 2080$		$\sigma_{I_{19}} = 2009$		$\sigma_{I_{18}} = 1956$		$\sigma_{I_{17}} = 1917$		$\sigma_{I_{16}} = 1904$		$\sigma_{I_{15}} = 1896$		$\sigma_{I_{14}} = 1880$		Признаки
	$\sigma_{I_{19}}$	$\sigma_{I_{19}} - \sigma_{I_{20}}$	$\sigma_{I_{18}}$	$\sigma_{I_{18}} - \sigma_{I_{19}}$	$\sigma_{I_{17}}$	$\sigma_{I_{17}} - \sigma_{I_{18}}$	$\sigma_{I_{16}}$	$\sigma_{I_{16}} - \sigma_{I_{17}}$	$\sigma_{I_{15}}$	$\sigma_{I_{15}} - \sigma_{I_{16}}$	$\sigma_{I_{14}}$	$\sigma_{I_{14}} - \sigma_{I_{15}}$	$\sigma_{I_{13}}$	$\sigma_{I_{13}} - \sigma_{I_{14}}$	
<i>a</i>	.2106	.0026	.2035	.0026	.1969	.0013	.1948	.0031	.1938	.0034	.1947	.0051	.1927	.0047	<i>a</i>
<i>b</i>	.2115	.0035	.2046	.0037	.1999	.0043	.1971	.0054	.1952	.0048	.1952	.0056	.1927	.0047	<i>b</i>
<i>c</i>	.2152	.0072	.2070	.0061	.2011	.0055	.1967	.0050	.1944	.0040	.1951	.0055	.1929	.0049	<i>c</i>
<i>d</i>	.2106	.0026	.2029	.0020	.1966	.0010	.1930	.0013	.1904	.0000	.1914	.0018	.1894	.0014	<i>d</i>
<i>e</i>	.2166	.0086	.2080	.0071	.2018	.0062	.1979	.0062	.1961	.0057	.1977	.0081	.1967	.0087	<i>e</i>
<i>f</i>	.2150	.0070	.2061	.0052	.1986	.0030	.1946	.0029	.1924	.0020	.1931	.0035	.1892	.0012	<i>f</i>
<i>g</i>	.2115	.0035	.2032	.0023	.1962	.0006	.1919	.0002	.1896	— .0008	—	—	—	—	<i>g</i>
<i>h</i>	.2123	.0043	.2036	.0027	.1972	.0016	.1948	.0031	.1912	.0008	.1922	.0026	.1911	.0031	<i>h</i>
<i>i</i>	.2145	.0065	.2069	.0060	.2022	.0066	.1978	.0061	.1960	.0056	.1948	.0052	.1962	.0082	<i>i</i>
<i>j</i>	.2123	.0043	.2042	.0033	.2055	.0099	.1956	.0039	.1963	.0059	.1954	.0058	.1992	.0112	<i>j</i>
<i>k</i>	.2026	— .0054	.1966	— .0043	.1942	— .0014	.1928	.0011	.1922	.0018	.1916	.0020	.1897	.0017	<i>k</i>
<i>l</i>	.2013	— .0067	.1956	— .0053	—	—	—	—	—	—	—	—	—	—	<i>l</i>
<i>m</i>	.2009	— .0071	—	—	—	—	—	—	—	—	—	—	—	—	<i>m</i>
<i>n</i>	.2018	— .0062	.1961	— .0048	.1917	— .0039	—	—	—	—	—	—	—	—	<i>n</i>
<i>o</i>	.2193	.0113	.2112	.0103	.2063	.0107	.2033	.0116	.1999	.0095	.2004	.0108	.2004	.0124	<i>o</i>
<i>p</i>	.2155	.0075	.2080	.0071	.2025	.0069	.1976	.0059	.1947	.0043	.1947	.0051	.1936	.0056	<i>p</i>
<i>q</i>	.2071	— .0009	.1992	— .0017	.1926	— .0030	.1904	— .0013	.1905	.0001	.1907	.0011	.1916	.0036	<i>q</i>
<i>r</i>	.2098	.0018	.2015	.0006	.1953	— .0003	.1917	.0000	.1905	.0001	.1907	.0011	.1916	.0036	<i>r</i>
<i>s</i>	.2100	.0020	.2016	.0007	.1956	.0000	.1915	— .0002	.1897	— .0007	.1880	.0016	—	—	<i>s</i>
<i>S</i>	.2123	.0043	.2052	.0043	.1983	.0027	.1932	.0015	.1916	.0012	.1898	.0002	.1928	.0048	<i>S</i>

признак. Вообще лишь в редких случаях признак сохраняет свое значение, если меняется комбинация признаков. Мы можем легко убедиться в этом, сравнивая различные графы табл. 8.

Особый интерес представляет окончательная комбинация признаков (графа седьмая). Здесь нет ни одного лишнего или вредного признака, вследствие чего определение таксономической ценности признаков является более строгим. Базируясь на величине разностей  $\sigma_{I_{13}} - \sigma_{I_{14}}$ , мы можем расположить 14 признаков в порядке возрастания их таксономического значения:

*o, j, e, i, p, c, S, a, b, r, h, k, d, f.*

В этом ряду ценность первого члена *o* превышает ценность последнего *f* более, чем в 10 раз. Отметим кстати, что первый признак очень слабо связан со всеми остальными. Благодаря такой независимости таксономическое значение *o* меняется лишь слабо при различных комбинациях признаков.

Чем объяснить тот факт, что *g* и *s*, казавшиеся полезными для характеристики при первоначальных комбинациях признаков, затем оказались вредными и были исключены? Нужно думать, что их влияние на величину  $\sigma_I$  компенсировалось влиянием других признаков. Мы видели, что отрицательная роль *s* обнаружилась лишь после выключения *l, m, n*, а аналогичная роль признака *g* — после исключения еще и *q*.

Возникает очень важный вопрос: можем ли мы поручиться в том, что наша окончательная комбинация из 14 признаков действительно является наилучшей для характеристики? Нельзя ли, составив иную комбинацию, добиться еще меньшей величины  $\sigma_I$ ?

Чтобы решить этот вопрос эмпирическим путем, нам пришлось бы перебрать все комбинации, которые можно составить из наших 20 признаков. Эта задача наталкивается на очень большие вычислительные трудности. Даже и те вычисления, которыми мы ограничились, потребовали очень большого труда. Тем не менее мы попробовали составить некоторые из этих комбинаций, именно, взяв те, которые по характеру коррелятивных связей казались нам сулящими некоторый успех. Однако все эти попытки дали отрицательный результат. Ясно, что таксономическая ценность той или иной комбинации признаков в конечном счете зависит от характера их коррелятивных связей, являясь некоторой функцией от коэффициентов корреляции *r*. Нахождение этой функции было нам, однако, не по силам.

Интересно сопоставить полученную нами величину  $\sigma_I = .1880$  с данными табл. 1. В последнем случае мы брали признаки более или менее наудачу, и поэтому даже 26 признаков дали худший результат, чем избранные 14 ( $\sigma_{I_{26}} = .1901$ ).

В качестве второго примера мы возьмем характеристику конечностей.

В табл. 9 приведены данные, относящиеся к передней конечности. Все 7 (индексных) признаков дают  $\sigma_{I_7} = 3233$ . Результат последовательного выключения признаков по одному приведен в первой графе таблицы. Мы видим, что роль признаков в уменьшении  $\sigma_{I_7}$  резко различна и много больше, чем в примере с пропорциями крыла. Прежде всего нужно отметить, что  $t_4$  не уменьшает, а увеличивает  $\sigma_I$ . Другие признаки уменьшают  $\sigma_I$  на очень различные величины: *t* — на 0359, а *F* — только на 0005, т. е. в 72 раза

меньше! Поставив признаки в порядке убывания их таксономического значения, мы получаем следующий ряд:

$$t_5, t_1, T, t_3, t_2, F, t_4$$

После выключения «вредного» признака  $t_4$  все прочие оказывают на  $\sigma_{I_6}$  уменьшающее действие (графа вторая). Из сравнения цифр обеих граф мы видим, что таксономическое значение всех признаков, кроме  $T$ , возросло. Ряд, построенный в порядке убывающего значения ( $t_5, t_1, t_3, t_2, T, F$ ), отличается от предыдущего только тем, что  $T$  переместилось с третьего места на пятое.

Таблица 9. Таксономическое значение признаков передней конечности. Признаки типа  $F:L$

Признаки	I		II	
	$\sigma_{I_7} = .3233$		$\sigma_{I_6} = .3205$	
	$\sigma_{I_6}$	$\sigma_{I_6} - \sigma_{I_7}$	$\sigma_{I_5}$	$\sigma_{I_5} - \sigma_{I_6}$
$F$	3238	0005	3236	• 0031
$T$	3323	0090	3279	0074
$t_1$	3469	0236	3471	0266
$t_2$	3284	0051	3284	0079
$t_3$	3309	0076	3302	0097
$t_4$	3205	— 0028		
$t_5$	3592	• 0359	3598	0393

Удельный вес признаков в характеристике попрежнему очень различен, хотя и не в такой степени, как при наличии  $t_4$ .

Итак, в данном случае при наличии всего лишь 7 признаков один из них оказался не просто излишним, но и искажающим характеристику.

Таблица 10. Таксономическое значение признаков средней конечности. Признаки типа  $F:L$ .

Признаки	I		II		III	
	$\sigma_{I_7} = .3414$		$\sigma_{I_6} = .3280$		$\sigma_{I_5} = .3258$	
	$\sigma_{I_6}$	$\sigma_{I_6} - \sigma_{I_7}$	$\sigma_{I_5}$	$\sigma_{I_5} - \sigma_{I_6}$	$\sigma_{I_4}$	$\sigma_{I_4} - \sigma_{I_5}$
$F$	3644	0230	3552	• 0272	3514	0256
$T$	3280	— 0134				
$t_1$	3765	0351	3704	0424	3705	0447
$t_2$	3348	— 0066	3278	— 0022		
$t_3$	3549	0135	3435	0155	3490	0232
$t_4$	3507	0093	3354	0074	3451	0193
$t_5$	• 3660	0246	3584	0304	3606	0348

Средняя конечность (табл. 10) дает еще более интересный результат: здесь 2 признака ( $T, t_2$ ) вызывают увеличение  $\sigma_{I_7}$ , притом первый из них — значительное (графа первая). Выключение  $T$  сильно меняет соотношение признаков (графа вторая), но «вредное» действие  $t_2$  остается, хотя и в ослабленной степени. Избавившись от этого признака, мы получаем окончательный результат (графа третья): все признаки уменьшают величину  $\sigma_I$ . Степень уменьшения неодинакова, но различия здесь не столь велики, как в случае передней конечности. Порядок убывающего таксономического значения таков:

$$t_1, t_5, F, t_3, t_4.$$

Мы выключали сначала  $T$ , затем  $t_2$ . Спрашивается, не станет ли  $T$  в отсутствие  $t_2$  более «полезным» признаком? Мы попробовали заменить им признак  $t_1$ , вычислив  $\sigma_I$  для комбинации  $F, T, t_3, t_4, t_5$ . В этом случае  $\sigma_{I_5} = .3720$ , т. е. она гораздо больше, чем  $\sigma_{I_6}$  при комбинации  $F, t_1, t_3, t_4, t_5$ .

Таблица 11. Таксономическое значение признаков задней конечности. Признаки типа  $F:L$

Признаки	I		II	
	$\sigma_{I_7} = .3386$		$\sigma_{I_6} = .3248$	
	$\sigma_{I_6}$	$\sigma_{I_6} - \sigma_{I_7}$	$\sigma_{I_5}$	$\sigma_{I_5} - \sigma_{I_6}$
$F$	.3248	— .0138		
$T$	.3569	.0183	.3410	.0162
$t_1$	.3432	.0046	.3422	.0174
$t_2$	.3531	.0145	.3415	.0167
$t_3$	.3459	.0073	.3317	.0069
$t_4$	.3596	.0210	.3423	.0175
$t_5$	.3553	.0167	.3426	.0178

Рассмотрим, наконец, заднюю конечность (табл. 11). Здесь один признак ( $F$ ) увеличивает  $\sigma_I$  (графа первая). После его выключения все признаки оказываются уменьшающими  $\sigma_I$ . При этом роль 5 признаков ( $T, t_1, t_2, t_4, t_5$ ) приблизительно одинакова — разность  $\sigma_{I_5} - \sigma_{I_6}$  колеблется здесь в узких пределах, от .0162 до .0174. Значение же признака  $t_3$  значительно ниже. Располагая признаки в порядке убывания их таксономической ценности, мы получаем следующий ряд:

$$t_5, t_4, t_1, t_2, T, t_3.$$

Сравнивая результаты, полученные для разных конечностей, мы замечаем, что таксономическое значение признаков меняется от случая к случаю. Во всех трех случаях нам пришлось выключать разные признаки. В среднем же наибольшее значение имеет признак  $t_5$ , занимающий первое или второе место в ряду. За ним следует  $t_1$ . Остальные признаки имеют гораздо меньшее значение, причем на последнем месте стоит  $T$ .

Итак, наш второй пример подтверждает первый. Неравноценность признаков и относительность их таксономического значения выступают здесь столь же резко, если не более, чем в первом примере.



Анализ таксономической ценности признаков, который мы выполнили в предыдущих главах, позволяет рассмотреть проблему исчерпывающей характеристики индивида на более высоком уровне.

Разбирая какой-нибудь орган с таксономической точки зрения и используя ряд признаков одной категории (пропорции крыла, конечностей), мы убеждаемся в том, что уже некоторая часть этих признаков позволяет исчерпать все элементы независимости. Даже при очень небольшом числе их остальные либо ничего не прибавляют к точности характеристики, либо уменьшают ее, искажая правильное представление о существующем сходстве индивидов.

Но ведь и в пределах одного органа возможны разные категории признаков. Как мы видели в третьей главе, привлечение большего числа признаков, характеризующих тоже пропорции крыла, дает значительно меньшую величину  $\sigma_I$ , чем наши избранные 14 признаков. Правда, мы получили, в конце концов, остановку уменьшения  $\sigma_I$ . Но в пределах того же органа мы могли бы взять признаки иного порядка, например, толщину крыловых жилок. Эти признаки, разумеется, будут очень сильно связаны с размером крыла, но в то же время они в какой-то мере будут и независимы от него. Нужно думать, что привлечение новой категории признаков позволит добиться хотя бы некоторого уменьшения  $\sigma_I$ , а это будет равносильно более совершенной характеристике.

Характеризуя индивидуальные особенности конечностей, мы можем привлечь такую группу признаков, как величину, количество и расположение щетинок (хет). По всей вероятности, эти признаки будут лишь частично связаны с пропорциями частей ноги и, таким образом, увеличат точность таксономической характеристики. Наконец, и окраска конечностей дает некоторые индивидуальные вариации.

Таким образом, считая полученную нами характеристику крыла исчерпывающей, мы должны иметь в виду, что исчерпаны лишь особенности пропорций жилкования крыла.

Наша главная задача заключается сейчас не в составлении исчерпывающей характеристики какого-либо органа, а в том, чтобы доказать ее возможность. Очевидно, что категории признаков, которые мы можем использовать, ограничены в своем числе, не говоря уже о том, что признаки разных категорий в той или иной степени связаны друг с другом. Мы не можем найти в пределах органа бесконечно большое число категорий признаков.

Учитывая, что число признаков каждой категории, нужных для исчерпывающей характеристики, является ограниченным и что количество таких групп тоже конечно, мы получаем сумму конечных чисел. Следовательно, для исчерпывающей характеристики органа по всем категориям признаков нам нет нужды обращаться к бесконечности. Но это и означает реальную возможность такой характеристики.

До сих пор мы говорили о характеристике отдельного органа. Однако индивидуальные особенности не обязаны отразиться полностью в каком-либо одном органе. Учитывая с целью исчерпывающей характеристики индивида все его органы, не рискуем ли мы все же столкнуться с бесконечностью?

Это опасение неосновательно. Число органов конечно. В данном случае исключается произвол, который имеет место при выборе отдельных признаков органа: здесь мы ограничены определенным числом конструктивных элементов организма. Ввиду того, что число органов конечно, а количество таксономически «полезных» признаков каждого из них тоже ограничено, мы опять получаем сумму конечных чисел и снова избегаем бесконечности.

Рассматривая индивидуальный организм как сочетание известного количества органов и сравнивая его с другими индивидами той же группы, мы встаем перед дилеммой: являются ли органы, взятые в целом, коррелятивно-связанными или же они независимы друг от друга? Естественно, что и здесь нас интересует прежде всего таксономическая сторона вопроса. В случае полной независимости органов мы должны были бы при таксономической характеристике индивидов подбирать признаки так, чтобы они представляли органы в общей характеристике. Другими словами, мы имели бы такой же случай, как тот, когда орган состоит из некоторого числа зон, независимых друг от друга.

Если же органы связаны, то вопрос о представительстве их теряет свое значение.

Как решить эту дилемму? Существующие методы установления корреляции сводятся к тому, что мы ищем связь между двумя признаками или же (множественная корреляция) рассматриваем, какие значения принимает  $(n-1)$ -й признак при различных комбинациях значений группы из  $n$  признаков. Конечно, мы можем брать попарно признаки других различных органов и таким образом судить о степени связанности последних. Однако это сравнение будет столь же произвольным, сколь и выбор пар признаков.

На основе принципа исчерпывающей характеристики мы можем получить объективное решение данного вопроса. Для этого нам нужно иметь исчерпывающие характеристики обоих сравниваемых органов, следовательно, определить набор признаков, таксономически ценных. Взяв далее характеристики наших  $N$  индивидов по обоим органам ( $I_1, I_2, I_3, \dots, I_n$  и  $I'_1, I'_2, I'_3, \dots, I'_n$ ), мы получим два ряда вариантов, которые и будем сравнивать почленно, т. е.  $I_1$  с  $I'_1$ ,  $I_2$  с  $I'_2$ , и т. д. Коэффициент корреляции определится совершенно так же, как и при наличии вариантов, взятых по двум обычным признакам.

Достаточно большое значение вычисленного таким способом коэффициента  $r$  покажет нам, что атипическое строение одного органа связано с атипическим же строением другого. Отсутствие же связи будет означать, что эти органы, взятые как целое, в своем изменении от индивида к индивиду дают совершенно случайные комбинации.

Мы выполнили несколько таких вычислений для нашей пробной группы *D. plumipes*. В качестве вариантов первого признака мы взяли ряд  $I_{14}$ , полученный нами ранее на основе использования 14 избранных признаков крыла, разобранных в предыдущей главе. Вариантами второго признака у нас были  $I_6$ , вычисленные по 6 избранным признакам передней конечности.

Корреляция обоих признаков оказалась практически равной нулю:  $r = .036 \pm .100$ . Этот результат показывает, что пропорции крыла изменяются независимо от пропорций передних конечностей. Атипическое строение крыла индивидов нашей группы может одинаково часто сочетаться с типичным и с атипичным строением этих конечностей.

К иному выводу нас приводит сравнение конечностей между собой, выполненное аналогичным образом.

Корреляция передних и средних конечностей выражается величиной  $r = .362 \pm .087$ , а корреляция средних и задних —  $r = .331 \pm .089$ .

Таким образом, между конечностями, взятыми как целое (точнее, между их пропорциями), существует вполне реальная, хотя и невысокая, связь: отклонение от типа в строении одной конечности влечет за собой до некоторой степени аналогичное отклонение и другой.

Полученные результаты не позволяют нам принять независимость органов как общий случай. Но это означает также, что при вычислении исчерпывающей характеристики не приходится составлять ее из признаков, равномерно представляющих каждый орган. Таксономическая характеристика должна опираться на признаки индивида, независимо от органов, а изменение  $\sigma_I$  покажет, какие из них являются излишними, какие — «вредными»

По поводу нашего утверждения, что число органов конечно и устанавливается вполне объективно, можно заметить следующее: не всегда понятие органа как составной части организма является определенным. Мы можем произвольно сужать и расширять это понятие: то, что один считает органом, другой может рассматривать как совокупность органов.

На это можно возразить, что такое дробление понятия всегда имеет границу. С другой стороны, если брать органы более дробно, то и корреляция между ними окажется сильнее, а это означает более быстрое исчерпание элементов независимости. Следовательно, таксономическая проблема от этого не усложняется.

Несколько лет назад я предложил понятие признака-индикатора (7), которое затем было принято Терентьевым в его работах (13—15). В основе понятия лежит та мысль, что некоторые признаки имеют максимальное таксономическое значение и поэтому могут заменять целые группы признаков малоценных. В то время я понимал это следующим образом.

Представим себе группу коррелятивно связанных признаков. Каждый из этих  $n$  признаков характеризуется неодинаковой степенью связи с остальными ( $n - 1$ ). Имея все попарные коэффициенты корреляции, мы можем найти для каждого признака  $j$  среднюю тесноту связи его с остальными, беря в качестве показателя среднюю величину ( $M_j$ ) из  $r_{ja}, r_{jb}, r_{jc}, \dots$ . Признаком-индикатором будет тот, который характеризуется наибольшим значением  $M_j$ , ибо он с наибольшей точностью позволяет восстановить размер отклонения каждого из остальных признаков от соответствующей средней. Таксономическая ценность такого индикатора очевидна.

От самой идеи индикатора мы не отказываемся и теперь, ибо таксономическая проблема действительно заключается в «индикации»: от бесконечно большого числа возможных признаков нам нужно перейти к ограниченному их количеству, которое вполне заменило бы все остальные признаки. Однако предложенное нами ранее определение и использование индикатора мы теперь считаем несовершенным. В самом деле, в основе предлагавшейся нами тогда операции лежит опять-таки концепция зональности и представительства этих зон в характеристике. Принятие зон наталкивает нас, как было показано выше, на затруднения, вытекающие из фактического распределения коррелятивных связей. Поскольку мы не можем принять независимость зон и даже самих органов, нам приходится констатировать и исчерпывать элементы независимости другим путем, путем применения критерия  $\sigma_I$ . Индикаторами будут как раз те признаки, которые максимально уменьшают величину  $\sigma_I$ , т. е. по возможности не зависящие друг от друга. В отдельных же случаях, когда зональность достаточно сильно выражена, совокупность признаков, которые представляют зоны, может дать сравнительно хороший результат. Однако полученная таким путем характеристика не будет совершенной.

Все сказанное относится к проблеме индикаторов внутривидового порядка, т. е. к характеристике особей. Позднее (9) мы выдвинули аналогичную проблему для высших категорий, после чего она в интересной форме была разработана Беклемишевым (1). Поскольку

в настоящей работе мы рассматриваем лишь строение вида, мы воздержимся пока от анализа этого важного вопроса.

В заключение мы должны вернуться к работе Терентьева (14), о которой уже упоминалось выше. Автор ее высказывает ряд здравых мыслей о задачах систематики, как самостоятельной дисциплины, а в специальной части задается целью разрешить вопрос о выборе таксономически ценных признаков.

Справедливо полагая, что суть вопроса заключается в анализе связи признаков, Терентьев подробно разбирает закономерности распределения корреляций у *R. ridibunda*. В основу своего анализа он кладет наше понятие признака-индикатора—в более разработанной и усовершенствованной форме.

Концепция автора такова. Истинными таксономическими признаками являются лишь индикаторы, набор которых составляет «таксономический коэффициент». Как мы уже говорили в 6-й главе, эти признаки служат у него представителями корреляционных плеяд. Метод парциальной корреляции используется автором для обнаружения этих плеяд. Например, исключая признак длины тела (индикатор), автор выделяет плеяду из 10 скелетных» признаков. Для этого десятка вновь находится индикатор, по исключению которого путем применения парциальной корреляции плеяда расчленяется на две меньшего объема («Kopf»-Plejade из 4 признаков и «Extremitäten»-Plejade из 3 признаков; прочие 2 признака теряют свою связь с остальными). Автор не производит дальнейшего расчленения плеяд, повидимому, в силу исчерпания запаса признаков. Однако логически ясно, что эта операция может осуществляться и далее и что всегда найдутся признаки, связанные друг с другом сильнее, чем с прочими.

Терентьев не указывает, каким образом его анализ корреляций и расчленение плеяд должно использоваться для таксономической характеристики. Если для нее использовать только признаки-индикаторы, то что же делать с теми признаками, которые по исключении индикаторов оказываются несвязанными? Пренебречь ими означало бы явную ошибку. Если же их учитывать, то по мере дробления плеяд к ним должны присоединяться все признаки, кроме индикаторов. При неограниченном возрастании числа признаков их число должно быть бесконечно велико. Но при этом условии и количество индикаторов тоже будет стремиться к бесконечности. Следовательно, мы будем лишены возможности дать исчерпывающую характеристику, и основная проблема останется нерешенной.

Подчеркивая здесь интерес работы Терентьева, мы вправе ждать от автора разъяснения всех этих вопросов.

## 9. КОНСТРУКЦИЯ ВИДА

Результаты, полученные нами в предыдущих главах, позволяют решить вопрос, поставленный в заглавии,—о конструкции вида с таксономической точки зрения. Как мы видели, эта проблема сводится к установлению закономерностей сходственных отношений, существующих внутри вида, а сходство мы условились определять в единицах видового типа, посредством величины  $I$ .

Допустим, что числа  $I$  определены так, что они представляют собой исчерпывающие характеристики особей нашего вида. Наша задача будет состоять в том, чтобы найти функцию распределения  $I$ , — иными словами, написать уравнение вида.

Можно было бы найти видовое уравнение следующим образом. Если нам известны средние квадратические отклонения  $n$  признаков, а также связывающие их корреляции (попарные коэффициенты  $r$  и

вычисленные на основании их коэффициенты множественной корреляции  $R$ ), то мы можем найти вероятность любой системы отклонений  $n$  признаков от соответствующих средних арифметических. Эта функция будет поверхностью  $n$  измерений (по числу взятых признаков), нормальной или уклоняющейся от нормальной, смотря по характеру распределения отдельных признаков и их связей. Поскольку каждый индивид можно рассматривать как систему коррелятивно связанных отклонений от видового типа, наша функция позволяет определить вероятность любого индивида. Но на основании знания вероятностей можно сказать, каков будет состав вида: сколько индивидов того или иного типа мы можем ожидать во взятой нами пробной группе.

К этому способу нужно сделать следующее критическое замечание. Применяя его, мы произвольно выбираем признаки, не считаясь совсем или считаясь недостаточно строго с тем, насколько избранные признаки действительно способны характеризовать отдельные индивиды. Выход из

этого положения совершенно очевиден. Нам нужно взять такой набор признаков, который способен дать исчерпывающую характеристику индивидов. В этом случае наша функция распределения дает действительно строгое решение вопроса о составе или конструкции вида.

Видовое уравнение такого типа в полной мере учитывает все избранные признаки, не лишая их индивидуальности. Кроме того, оно учитывает и знак отклонения в каждом отдельном признаке. Не умаляя ценности такого решения вопроса, мы

можем иметь в виду и другое, — в форме закона распределения чисел  $I$ .

В этом случае мы отвлекаемся от отдельных признаков и от знака отклонений — речь идет лишь о средней удаленности индивидов от видового типа, и наша функция должна дать вероятность той или иной величины  $I$ .

Каково же в действительности распределение чисел  $I$ ? Мы можем решить этот вопрос эмпирически, используя наш материал по *Dolichopus plumipes*. В качестве примера мы возьмем распределение  $I_{14}$  — чисел, которые представляют собой характеристики индивидов нашей группы по 14 избранным признакам крыловых пропорций.

Мы видели, что среднее квадратическое отклонение этого ряда ( $\sigma_{I_{14}}$ ) равно 0.1880. Величина вариантов колеблется от 0.38 до 1.39.

На рис. 4 приведены полигон частот ряда и соответствующая  $\sigma = 0.1880$  нормальная кривая распределения.

На первый взгляд эмпирическая линия кажется эксцессивной. Однако вычисление эксцесса дает незначительную по сравнению с ошибкой величину:

$$ex = .2704 \pm .4899.$$

Несколько сильнее выражена асимметрия:

$$as = .3456 \pm .2449.$$

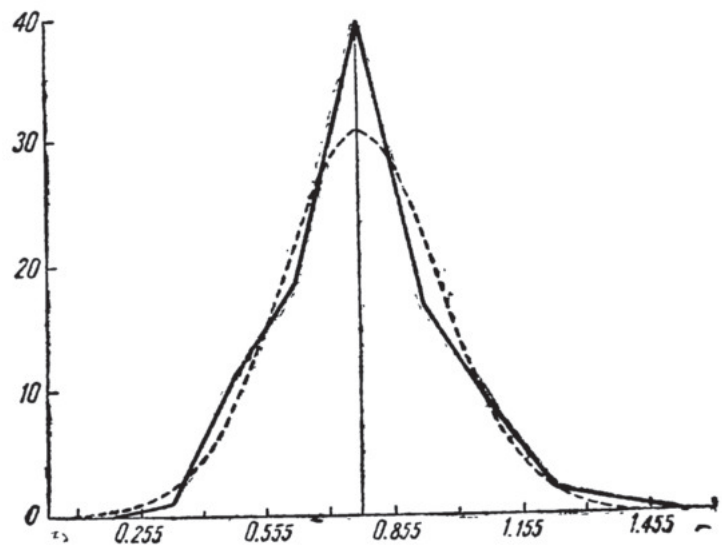


Рис. 4

В общем же наша эмпирическая линия не отличается сколько-нибудь значительно от нормальной кривой.

Итак, наибольшую часть пробной группы составляют особи, среднеудаленные от типа, или умеренно типичные. Меньше всего особей резко атипичных или, наоборот, сильно приближающихся к типу.

Такая картина получается благодаря тому, что мы брали абсолютные величины отклонений особей в отдельных признаках. Если бы мы учитывали алгебраические суммы, то происходило бы взаимное уничтожение положительных и отрицательных отклонений, и максимальное количество особей группировалось бы около видового типа. Однако, если мы ставим вопрос о точном определении сходства особей с типом, нам становится безразличным, в каком направлении отклоняются особи: нас интересует лишь степень отклонения, степень типичности.

Аналогичный результат получился и для распределения  $I$ , вычисленных по избранным (см. выше) признакам пропорций конечностей.

Характеристики, выражаемые числами  $I$ , хотя и являются исчерпывающими, но лишь для определенных категорий признаков некоторых органов. Спрашивается, каково же распределение чисел  $I$ , найденных как истинные исчерпывающие характеристики,—с учетом всех таксономически ценных признаков.

Нами было достаточно показано в предыдущих главах, что такие характеристики должны оставить в силе индивидуальные различия в степени типичности. Мы доказали, что закономерность, утверждавшаяся Heincke,—равная типичность индивидов одной расы—не имеет места в действительности, поскольку существуют корреляции, а число независимых признаков ограничено. Сохраняющаяся при любом числе таксономически ценных признаков изменчивость особей, очевидно, будет тоже подчиняться закону нормального распределения. (Дело в принципе не изменится, если мы в отдельных случаях будем получать какую-либо вариацию нормальной кривой Гаусса.)

Этот вывод с необходимостью вытекает из того факта, что в пределах особи отклонения комбинируются до известной степени случайно. Элементы независимости признаков, которые мы стремимся исчерпать в наших характеристиках, являются порукой в том, что и распределение индивидов по степени типичности будет в значительной степени регулироваться законом Гаусса.

Наш анализ таксономических признаков необходимо дополнить еще следующими соображениями.

Составляя характеристику или описание особи, мы принуждены опираться на какие-то отдельные свойства ее морфологических и функциональных компонентов или на некоторые элементы отношения особи к окружающей ее обстановке, физической и биотической. Это и есть признаки. Было бы нелепостью считать, что индивид «состоит» из признаков. Напротив, признаки надо рассматривать как различные проявления индивидуальной сущности.

Поскольку же выбор признаков всегда в известной мере произволен, мы не можем избежать в своих описаниях элемента субъективности, а это на первый взгляд ставит под знак вопроса ценность составляемых характеристик. Но эта опасность—кажушаяся. Доказанная нами возможность исчерпывающей характеристики особи полностью ее аннулирует. Дело не в том, какие признаки мы избираем, главное—как их брать. Основная задача есть исчерпание всех элементов независимости при помощи критерия  $\sigma_I$ . Мы можем достичь этой цели разными средствами, используя различные признаки. Нетрудно представить себе две группы признаков одинакового

таксономического значения: каждая из них дает исчерпывающую характеристику особи. Таким образом, несмотря на субъективизм в выборе признаков, возможно вполне объективное решение проблемы.

## 10. ЗАКЛЮЧЕНИЕ

Принцип исчерпывающей характеристики дает возможность точного определения степени сходства. Благодаря этому мы можем дать строгую оценку строения систематических групп. Разумеется, наш принцип применим не только к элементарному виду или расе, но и к группам высшего порядка. Это применение требует, однако, дальнейшей разработки вопроса, так как, восходя от элементарных групп к высшим, мы сталкиваемся с новыми специфическими проблемами.

Достаточно указать на то, что вариантом высшей группы является уже не индивид, а некоторая средняя—видовой, родовой и т. д. тип. Затем в этих группах часто приходится иметь дело с качественными отличиями вариантов, которых нам не пришлось учитывать при анализе вида. Несмотря на это усложнение, принцип исчерпывающей характеристики вполне применим и здесь.

Познание закономерностей строения систематических категорий, помимо самодовлеющего значения, не может остаться без влияния для различных отраслей биологии, имеющих дело с систематическими единицами. Но ведь совершенно очевидно, что в них заинтересованы все биологические дисциплины.

Основа точного определения сходства есть правильный выбор признаков—вопрос, которому мы и посвятили большую часть нашего исследования. Без разрешения этой проблемы мы не можем ожидать прогресса систематики.

В недооценке этого обстоятельства мы видим основной недостаток предложенного К. Пирсоном «коэффициента расового сходства» (5). Этот коэффициент, позволяющий устанавливать степень сходства двух рас, выводится автором на основании предположения о независимости различных признаков друг от друга, а применяется без всякого учета существующих между ними корреляций как Пирсоном, так и его многочисленными последователями. В результате такого применения коэффициента получается искаженная картина сходства, и главное здесь то, что мы не можем судить о размерах этого искажения. Наш метод исчерпания элементов независимости свободен от такого недостатка.

Практика применения  $\sigma_1$ , конечно, требует разработки, на что мы уже указывали выше. Чисто эмпирическое испытание различных комбинаций признаков с точки зрения оценки величины  $\sigma_1$  наталкивается на очень большие технические трудности. Впрочем, они значительно меньше при исследовании систематических групп высшего порядка, что мы надеемся показать в следующей статье.

Принцип исчерпывающей характеристики на практике приводит нас к составлению совершенных диагнозов.

Всякая наука, имеющая надобность в систематизации объектов, образующих естественные группы, не может обойтись без помощи таких диагнозов. Таким образом, для нашего принципа открывается широкое поле применения.

## ЛИТЕРАТУРА

1. Beklemischev W., Zool. Jahrb., Abt. Syst., 54, 1927.—2. Heinicke F., Naturgeschichte des Herings, 1898.—3. Ильин Н., Тр. Лабор. эксп. биол. Моск. зоопарка, 4, 1928.—4. Любичев А., Изв. Биол. ин-та Пермск. гос. ун-ва, 2, 1923.—5. Pearson K., Biometrika, 18, 1926.—6. Смирнов Е., Русск. зоол. журн., 3, 1923.—

7. Он же, Докл. Росс. акад. наук, 1924.—8. Smirnov E., Zool. Anzeiger, 61, 1924.—  
9. Он же, Ztschr. ind. Abst.-lehre., 37, 1925.—10. Он же, Biol. Generalis, 2, 1926.—  
11. Он же, Ztschr. ind. Abst.-lehre, Suppl., 2, 1928.—12. Smirnov E. u. Zhelochovtsev A., Ztschr. wiss. Zool., 135, 1929.—13. Терентьев П., Русск. гидробиол. журнал, 7, 1928.—14. Terentjev P., Biometrika, 23, 1931.—15. Терентьев П., Изв. Акад. наук СССР, 1936.

## SPECIES CONSTRUCTION FROM A TAXONOMIC POINT OF VIEW

by Eugen Smirnov

Institute of Zoology, Moscow State University

The taxonomic problem of a species consists in finding those regularities which govern the resemblance of individuals. With this purpose in view, the author analyses by means of biometrical methods characters peculiar for the structure of the wing and legs of a dipterous insect, *Dolichopus plumipes*, Scop.

The greatest difficulty in solving the problem named arises from the presence of infinite number of characters, which naturally renders impossible an accurate evaluation of similarity among the individuals of the species. The author overcomes the above difficulty by using the principle of exhaustive characteristic advanced by him. The idea of the character exhaustion serves as a base of that principle.

The author demonstrates that the number of taxonomically valuable characters instead of being infinite is limited by a comparatively small quantity. In case of detection and utilisation of those characters (for which a special method is offered), all other properties of individuals will be either tautological or distorting the real aspect of similarity.

By using his method, the author establishes the fact that the distribution of individuals according to the degree of their conformity to the type submits to Gauss' law of normal distribution. At the same time the largest number of individuals is formed of those which belong to the medial type, while individuals, characterised by a maximal and minimal distance from the specific type are seen to occur in an insignificant number.

---