Future cars: The electric option

A. K. Shukla, N. R. Avery and B. C. Muddle*

Global economic, environmental and political issues are pushing car manufacturers to consider electric power systems as an alternative to the current spark ignition engine. In this article, we analyse the power and energy requirements of a modern car and conclude that a viable electric car could be operated with a 50 kW polymer electrolyte fuel cell (PEFC) stack to provide power for cruising and climbing, coupled in parallel with a 30 kW supercapacitor or battery bank to supply additional short term burst power during acceleration.

ENERGY is the pervasive element of a modern industrial economy. A substantial proportion of our present day energy need is met through fossil fuels derived from ultimately finite reserves and thus cannot be sustained indefinitely in the longer term. Besides, the deleterious effects of excessive consumption of carbonaceous fuels on the economy and ecology of a large part of the world is already apparent. While much debate surrounds the utilization of fully renewable energy sources like solar, hydro, biomass, etc. these will be unable to provide adequate energy into the foreseeable future. Instead, it is being increasingly realized that there is an urgent need to extract more useable energy from present non-renewable fuels.

Of all the sources of carbonaceous fuels, petroleum is by far the most convenient and therefore valuable. However, it is the least abundant and not widely distributed, rendering countries with the largest reserves disproportionate economic and political sway. Most recent estimates have suggested that, at present and projected discovery, production and consumption rates, world oil-supply will fail to meet demand by about the year 2010 (refs 1–4). Concerns of this kind have brought into sharp focus the need to develop new, more energy efficient and environmentally benign energy systems. Indeed, the development of systems of this kind is expected to become one of the significant drivers of national economies into the next millennium.

The vast majority of existing energy generating systems are of the thermo-mechanical type involving reciprocating engines or rotary turbines. They use the fuel in a controlled internal combustion or by raising high pressure steam. While well known reversible thermodynamic considerations impose a predictable efficiency limitation on heat engines of these kinds, additional irreversible losses associated with the expanding combustion products in the confinements of combustion chamber also occur. For example, with large multi-megawatt thermo-mechanical systems like diesel engines and gas turbines, thermal efficiencies in excess of 40% may be achieved and, with a combined cycle to raise steam from their hot exhaust, gas turbines can achieve values approaching 60%. By contrast, the smaller 100 kW class Otto cycle, spark-ignition engines as used in modern cars, can achieve little more than 20% efficiency under a typical range of driving conditions.

It has long been considered that electric traction for cars is a viable way of improving both the emission and energy efficiency of these vehicles. In the urban environment, battery-powered cars have the considerable advantages of essentially zero-emissions and high energy conversion efficiency. However, when overall emission and energy efficiency is extended to include battery charging and discharging, power generation and distribution, the actual global gains may not be so attractive[5]. Any actual gains will depend on the battery type and its operational characteristics as well as the energy source used in the power station. For example, it has been estimated that with the mixture of energy used to generate electrical power in Australia, battery-powered vehicles could lead to a 20% reduction in CO_2 emission[6].

In response to increasing concerns over urban air quality, the State of California enacted in 1994, legislation requiring that by 1998, 2% of cars offered for sale be zero-emission, increasing to 5% by 2000 and ultimately to 10% by 2003. These deadlines have subsequently been amended, largely because of the failure of battery-powered vehicles, which were originally seen as the solution, to perform at a level approaching that of the existing spark ignition-powered cars. Indeed, pure battery- powered cars are no longer regarded as an acceptable or viable alternative, except possibly for local commuter use. More recently, the advantage of on-board electricity generation in a fuel cell has been recognized and a considerable effort has been mounted internationally to develop systems with acceptable performance and cost to address the needs of the modern motorist.

A. K. Shukla is at the Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India; N. R. Avery is at Manufacturing Science and Technology, CSIRO, Private Bag 33, Clayton South MDC, Vic. 3169, Australia; and B. C. Muddle is at Department of Materials Engineering, Monash University, Clayton, Vic. 3168, Australia.
*For correspondence (e-mail: shukla@sscu.iisc.ernet.in)

Traction kinematics

To properly assess options for systems to power electric vehicles, it is necessary to estimate quantitatively the power and energy required to propel a modern car of the type shown in Figure 1. Neglecting relatively minor losses due to road camber and curvature, the power required at the drive wheel ($P_{traction}$) may be written as:

$$P_{traction} = P_{grade} + P_{accel} + P_{tyres} + P_{aero} \,. \quad (1)$$

The first two terms in eq. (1) describe the rates of change in potential (PE) and kinetic (KE) energies associated with climbing and acceleration, respectively. The power required for these actions may be estimated from Newtonian kinematics, thus:

$$P_{grade} = \mathrm{d(PE)/dt} = Mgv \sin \theta, \text{ and} \quad (2)$$

$$P_{accel} = \mathrm{d(KE)/dt} = Mv\, \mathrm{d}v/\mathrm{d}t = Mav, \quad (3)$$

where M is the mass of the car, v its velocity, a its acceleration and tan θ the gradient.

The potential and kinetic energies acquired by the car as a result of climbing and acceleration represent reversibly stored energy and, in principle, may be recovered by appropriate regenerative methods.

The last two terms in eq. (1) describe the power which is required to overcome tyre friction and aerodynamic drag which are irreversibly lost, mainly as heat and noise and cannot be recovered. The power required here may be estimated from the following empirical relations:

$$P_{tyres} = C_t Mgv, \quad (4)$$

and

$$P_{aero} = 0.5\, dC_a A(v + w)^2 v, \quad (5)$$

where C_t and C_a are dimensionless tyre friction and aerodynamic drag coefficients, respectively, d is the air density, w is the head wind velocity, g is gravitational acceleration, and A is the frontal cross-sectional area of the car.

From the parameters associated with a typical modern medium size car, viz. $M = 1400$ kg; $A = 2.2\ \mathrm{m}^2$; $C_t = 0.01$; $C_a = 0.3$; $d = 1.17\ \mathrm{kg/m}^3$; $g = 9.8\ \mathrm{m/s}^2$, its power requirements may be estimated from eqs (2)–(5). For the irreversible losses, eqs. (4) and (5) show that while P_{tyres} is linearly dependent on velocity, P_{aero} varies as the third power of velocity and although negligible at low velocity, the latter becomes the dominant irreversible loss at high speed. As an example, for these parameters, for a car traveling at about 50 km/h, tyre friction is twice the aerodynamic drag and together amount to about 3 kW. At 100 km/h highway cruising, aerodynamic drag increases considerably to over twice the tyre friction, increasing total power requirement to about 12 kW.

Taking the example of a hill with a substantial 10% gradient, climbing at 80 km/h requires about 38 kW, including tyre friction and aerodynamic drag. Acceleration is more demanding, particularly at high velocity. For example, acceleration at 5 km/h/s requires 29 kW at 50 km/h and increases to 66 kW at 100 km/h.

The above estimates are for the power supplied to the wheels of the car and do not include losses incurred in delivering that power to the wheels. At this time in the development of electric traction systems, a precise estimate of this is difficult to obtain. Anecdotal information suggests that the efficiency of the power conditioning electronics, together with the electrical and mechanical drive train (ε_{drive}), is likely to be about 0.85. Additional power (P_{access}) may also be required to power accessories like radios, lights, steering, air-conditioning, etc. which is likely to add about 4 kW to the total power demand of the car.

In this way, the instantaneous total power required from the electric power system (P_{total}) will be given by:

$$P_{total} = P_{traction}/\varepsilon_{drive} + P_{access} \,. \quad (6)$$

An analysis of this kind indicates that the power plant of a modern car must be capable of delivering about 50 kW of sustained power for accessories and hill climbing, with burst power for a few tens of seconds to about 80 kW during acceleration. For a car with these performance characteristics, this sets the upper power limit required, with more common usage rarely exceeding 15 kW while cruising.

Equations (2)–(5) show that with the exception of aerodynamic losses, the power requirements scale with the mass of the car. While new light-weight materials and construction methods, together with design innovations to reduce aerodynamic drag may be able to reduce the power and energy requirements of the car, gains in excess of about 25% will be difficult to achieve within an acceptable cost regime.

The energy consumed by a car is simply the time integral of the traction and accessory power plus that consumed by the power plant at idle (E_{idle}) thus:

$$E_{total} = \int P_{total}\, \mathrm{d}t + E_{idle} \,. \quad (7)$$

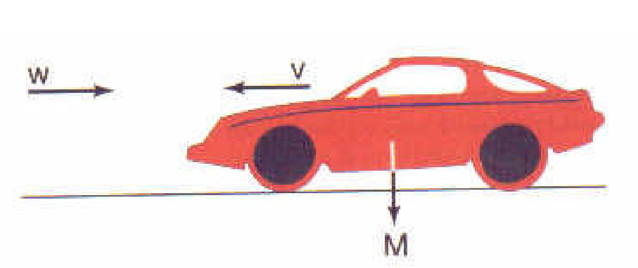

Figure 1. A modern spark-ignition engine car; M is its mass, v its velocity and w is the head-wind velocity.

E_{total} depends on the nature of the drive cycle the car is required to perform. Many estimates have been made for selection of acceleration – climb – cruise – idle sequences designed to simulate common driving practices. As with the power requirements of the car, E_{total} is very dependent on car mass. For the car parameters listed above, this may be expected to be near 200 Wh/km. For a modern spark ignition engine operating at about 20% efficiency under normal driving conditions, this corresponds to a rate of petrol consumption of about 10 litres/100 km.

From this quantitative analysis of the power and energy requirements of a car, the suitability of the various electric power options may be objectively evaluated.

Electric power options

In the preceding section, both the power and energy demands of a modern car were discussed. In brief, it was concluded that a base sustainable power of 50 kW, supplemented with short acceleration bursts to 80 kW will suffice in most driving requirements. Energy demand depends greatly on driving characteristics, but under normal usage can be expected to be about 200 Wh/km. The question remains as to which electric power option best satisfies these requirements. In addition, the system should be of robust and of compact construction and with near-ambient temperature operation to facilitate rapid start-up and intermittent usage. We will now consider the relative merits of supercapacitors, storage batteries and fuel cells to achieve these objectives.

Supercapacitors

The charge separation, which occurs at the interface between the electrode surface and the electrolyte acts as a pseudo-capacitor and as such is capable of storing significant amounts of energy. Such an electrolytic capacitor can store far greater amounts of energy than a simple capacitor consisting of two plates separated by gas or vacuum as dielectric. New supercapacitors based on this principle require conducting electrodes with surface areas in excess of 2000 m^2/g and are typically fabricated from activated carbons. Additional pseudo-capacitance at, for example, a highly-dispersed ruthenium oxide surface, can further enhance their energy capacity. Since the stored energy is essentially non-redox, supercapacitors may be rapidly charged and discharged, and are capable of delivering power densities in excess of 1 kW/kg, albiet at relatively low energy densities, typically less than 5 kWh/kg. Nevertheless, supercapacitors appear ideally suited for delivering the burst acceleration power requirements of an electric car at a relatively low weight penalty.

Storage batteries

Batteries store a fixed amount of chemical energy and may be recharged when the electrochemically active materials in them have been exhausted. As a result of a largely technically mature range of design, electrode type, electrolyte and operating temperature, modern rechargeable storage batteries offer a significant diversity of operational characteristics, including power and energy density, and life-time, balanced always by the cost of the system. Storage batteries which may be considered for electric car applications are given in Table 1 along with their energy and power densities.

From the cost perspective, lead-acid batteries appear the most attractive, but with an energy density of only 35 Wh/kg, almost 6 kg of battery is required to drive a car for 1 km. Coupled with its relatively slow recharge characteristics, it is immediately apparent that the lead-acid battery, in spite of its technical maturity and low cost, is an unacceptable option. Indeed, the failure of General Motors EV1 to find consumer acceptance may, at least in part, be linked to its dependence on lead-acid batteries[5]. When the high-temperature batteries are rejected for their inability to offer acceptable intermittent operational performance, it is seen from Table 1 that the most viable battery systems are either the nickel-metal hydride or lithium-ion types. Even neglecting the relatively high cost of these battery systems, their energy densities still require 2–2.5 kg of battery to travel 1 km. At 1 kg/km, the zinc–air battery is approaching an acceptable performance but technical difficulties in producing a truly rechargeable system continue to

Table 1. A comparison of the most promising storage batteries for electric cars

Cell type	Nominal (V)	Specific energy (Wh/kg)	Energy density (Wh/l)	Specific power (W/kg)	Power density (W/l)	Self discharge life (%/month)	Cycle	Comments
Lead–acid	2.0	35	70	~ 200	~ 400	4–8	250–500	Least-cost technology
Lithium–ion	3.6	115	260	200–250	400–500	5–10	500–1000	Intrinsically safe; contains no metallic lithium
Lithium–polymer	3.0	100–200	150–350	> 200	> 350	~ 1	200–1000	Not yet available commercially; contains metallic lithium
Nickel–cadmium	1.2	40–60	60–100	140–220	220–350	10–20	300–700	Exhibits memory effect and contains toxic cadmium
Nickel–metal hydride	1.2	60	220	130	475	30	300–500	No memory effect; cadmium free
Zinc–air	1.2	146	204	150	190	~ 5	~ 200	Requires air-management
$Na–NiCl_2$	2.6	100	160	150	250	~ 1	~ 1000	High-temperature

frustrate their implementation.

Fuel cells

For their high power and energy densities, fuel cells are emerging as the most viable candidate for powering electric cars. Like storage batteries, fuel cells deliver energy by consuming electroactive chemicals, but differ significantly in that these chemicals are delivered on-demand to the cell. As a result, a fuel cell can generate energy continuously and for as long as the electroactive chemicals are supplied to the cell. Typically, these chemicals consist of a fuel to the anode and air to the cathode. Although the first hydrogen–oxygen fuel cell was demonstrated as early as 1839 by Sir William Grove, most fuel cell developments have taken place in the last 50 years; initially with developments for the space industry of the sixties, followed by the energy crisis of the seventies and more recently, with the push to electric cars. With the profound advances in fuel cell technology that have been achieved in the last decade, there is emerging of a truly new electrochemical energy technology with the potential to challenge the *status quo* of combustion engines. Most recent developments in PEFC are delivering systems with power densities approaching 30 W/kg with overall energy efficiencies substantially greater than the 20% of a modern car spark ignition engine. While the future of the fuel cell to meet emerging energy needs is promising, the challenge today remains to reduce the cost of these systems to a level where they can properly assert their intrinsic efficiency and environmental advantages.

There are many types of fuel cells differentiated by the nature of the electrolyte as shown in Table 2. The suitability of a fuel cell for a given application depends mainly on its operational temperature.

Perspective for electric cars

The relative power and energy densities of current generation supercapacitors, storage batteries and fuel cells may be summarized in a Ragone diagram of the type shown in Figure 2. As discussed earlier, the overall energy demand of a modern electric car is about 200 Wh/km, in which case the power plant should be capable of delivering in excess of 500 Wh/kg to keep the power plant within acceptable weight limits. It is immediately apparent that this cannot be achieved by any of the direct electric storage systems, viz. supercapacitors or batteries, and that it is unlikely to be an achievable target for these systems within the foreseeable future. At best, storage battery and supercapacitor-powered vehicles may be able to establish a limited niche in short-range commuter vehicles, especially where the relatively long recharge times can be accommodated by the pattern of vehicle usage.

For transportation, the low operating temperature and rapid start start-up characteristics, together with its robust solid-state construction gives the polymer electrolyte fuel cell (PEFC) a clear advantage for application in cars. A typical construction of a PEFC is shown schematically in Figure 3 (ref. 7). Its energy conversion efficiency at the operating cell voltage of 0.6 V is about 50%, much higher than any spark ignition engine.

The preferred fuel for the PEFC is hydrogen. While many

Table 2. Contending fuel cell technologies and their operational temperatures

Type	Electrolyte	Fuel	Operating temperature (°C)		Cell reactions
Solid polymer electrolyte fuel cell (PEFC)	Solid–polymer electrolyte membrane (Nafion)	Hydrogen	60–90	Cathode : Anode : Cell reaction :	$O_2 + 4H^+ + 4e^- \quad 4H_2O$ $2H_2 \quad 4H^+ + 4e^-$ $O_2 + 2H_2 \quad 2\,H_2O$ (1.23 V)
Solid polymer electrolyte direct methanol fuel cell (SPE–DMFC)	Solid–polymer electrolyte membrane (Nafion)	Methanol	60–90	Cathode : Anode : Cell reaction :	$3/2\,O_2 + 6H^+ + 6e^- \quad 3H_2O$ $CH_3OH + H_2O \quad CO_2 + 6H^+ + 6e^-$ $3/2\,O_2 + CH_3OH \quad CO_2 + 2H_2O$ (~ 1.19 V)
Alkaline fuel cell (AFC)	Aqueous KOH	Hydrogen	60–120	Cathode : Anode : Cell reaction :	$O_2 + 2H_2O + 4e^- \quad 4OH^-$ $2H_2 + 4OH^- \quad 4H_2O + 4e^-$ $O_2 + 2H_2 \quad 2H_2O$ (1.23 V)
Phosphoric acid fuel cell (PAFC)	Phosphoric acid	Hydrogen	180–210	Cathode : Anode : Cell reaction :	$O_2 + 4\,H^+ + 4\,e^- \quad 2H_2O$ $2H_2 \quad 4H^+ + 4e^-$ $O_2 + 2H_2 \quad 2H_2O$ (1.23 V)
Molten carbonate fuel cell (MCFC)	Molten carbonate melts (Li_2CO_3/Na_2CO_3)	Hydrogen	600–700	Cathode : Anode : Cell reaction :	$1/2\,O_2 + CO_2 + 2e^- \quad CO_3^{2-}$ $H_2 + CO_3^{2-} \quad H_2O + CO_2 + 2e^-$ $H_2 + 1/2\,O_2 \quad H_2O$ (~ 1.5 V)
Solid oxide fuel cell (SOFC)	Yttria-stabilized zirconia ($ZrO_2–Y_2O_3$)	Methane/ hydrogen	1000	Cathode : Anode : Cell reaction : or	$O_2 + 4e^- \quad 2O^{2-}$ $2H_2 + 2O^{2-} \quad 2H_2O + 4e^-$ (hydrogen) or $1/2\,CH_4 + 2O^{2-} \quad H_2O + 1/2\,CO_2 + 4e^-$ (methane) $O_2 + 2\,H_2 \quad 2\,H_2O$ (~ 1.1 V) (hydrogen) $O_2 + 1/2\,CH_4 \quad H_2O + 1/2\,CO_2$ (~ 1.1 V) (methane)

Figure 3. Schematic representation of a polymer electrolyte fuel

strategies for providing hydrogen to the PEFC are presently being evaluated, the most acceptable proposal appears to be to generate it on-board and on-demand from liquid hydrocarbons or methanol. The technical challenge, however, is to modify large-scale industrial processes like steam reforming or partial-oxidation to lightweight reactors that can fit inside a car.

Prototype cars recently unveiled by Chrysler–Daimler, shown in Figure 4, are based on steam reforming of methanol. Methanol is relatively easy to process on-board and may be conveniently distributed through the existing service-station infrastructure. Meanwhile, partial oxidation reactors for processing gasoline, which is favoured by the Department of Energy in the US, are being developed by Epyx Corporation.

Partial oxidation offers compact reactors, fast start-up and rapid dynamic response while steam reforming produces more hydrogen with an increased fuel efficiency. Johnson Matthey in the UK have developed a methanol processor called the HotSpot which uses the heat produced by the partial oxidation process to drive the steam reforming reaction. This compact system ensures fast start-up and optimum fuel efficiency. The present generation HotSpot module can generate over 750 litres of hydrogen per hour, enough to power a 1 to 1.5 kW PEFC, depending on the operating voltage.

The hydrogen produced by steam reforming and partial oxidation of methanol (see Figure 5) contains carbon dioxide (20%) and a trace of carbon monoxide (2%). At the low operating temperature of the PEFC, carbon monoxide at levels in excess of about 0.001% poisons the platinum catalyst at the anode. Two strategies to circumvent this are under investigation. Either the carbon monoxide must be removed from the hydrogen stream in a separate process or new carbon monoxide-tolerant catalysts developed for deployment at the anode.

An elegant solution to the problems associated with the need

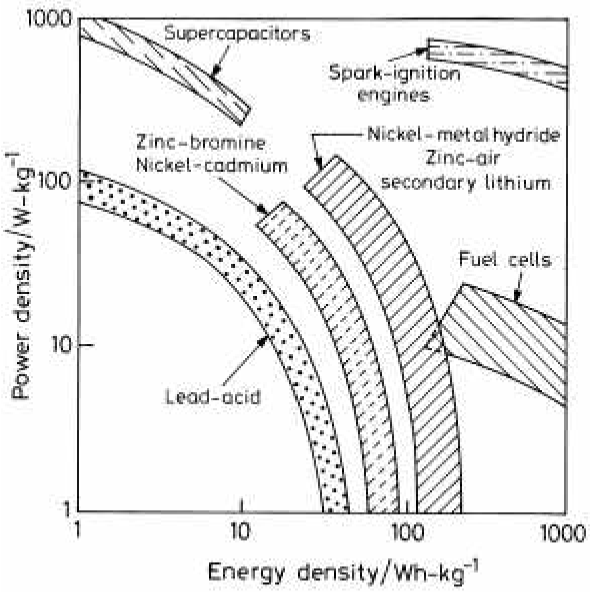

Figure 2. A Ragone plot comparison of power and energy densities of supercapacitors, storage batteries, fuel cells and spark ignition engines. In order to affect the appropriate comparisons, the energy densities of the spark ignition engines and fuel cells include typical

for gaseous hydrogen fuel, lies in operating the PEFC directly with a liquid fuel. Much consideration is therefore being given to PEFCs that run on air plus a mixture of methanol and water. The main technological challenges here are to develop better anode

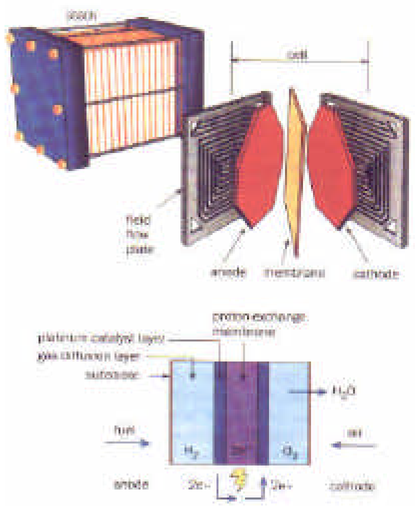

Figure 4. The Chrysler–Daimler prototype electric car.

catalysts to overcome efficiency losses at the anode and to improve the membrane electrolyte, and cathode catalysts to prevent methanol poisoning of the cathode.

It has been estimated that such a direct methanol fuel cell (DMFC) producing around 0.25 W/cm^2 of electrode, would be about the same size as a conventional methanol reformer/PEFC system operating at a power density of about 1 W/cm^2. Indeed, the development of a commercially feasible solid polymer electrolyte direct methanol fuel cell (SPE–DMFC) would be considerably simpler in both its construction and operation, and is widely regarded as the ‘holy grail’ of fuel cell technologies.

At present, the PEFC is emerging as the most viable electric power option for cars. Since the energy density of the PEFC power plant and fuel is similar to that of the present day spark ignition engines, comparable driving ranges may be expected. However, the power density of present PEFC systems tends to be less than that of spark ignition engines. Although the 80 kW of power needed to provide the acceleration could be supplied by a large PEFC alone, this will probably make the first generation system excessively large and heavy. Additionally, the high cost of newly developed fuel cells will persuade the car makers to use the smallest cells that will provide the required base power needs of about 50 kW. An acceptable compromise could be achieved with a supplementary parallel electric storage system using either high power density supercapacitors or less likely, storage batteries to provide the short duration acceleration. This electric storage system could also be used to regeneratively recover the energy which would be otherwise lost during braking. Since energy density is less important than power density for the acceleration of a car, supercapacitors would appear to be a superior alternative to any of the present storage battery options.

Figure 4. The Chrysler–Daimler prototype electric car.

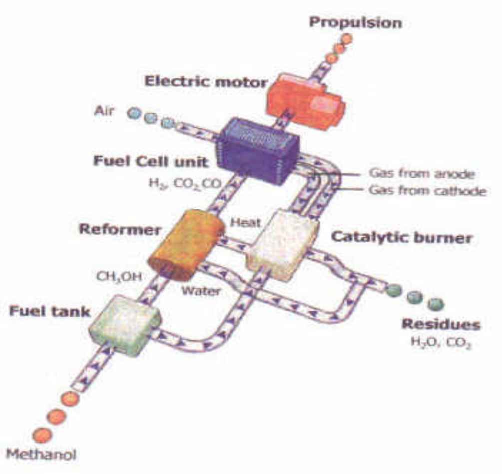

Figure 5. Schematic representation of the methanol reformer to be used for on-board reforming in the electric car.

Conclusion

From the foregoing, it is seen that a satisfactory system to meet the power and energy need of a modern car could be met with a 50 kW PEFC stack coupled in parallel with a 30 kW supercapacitor and/or battery bank to provide the short term power for acceleration. With the development of electric propulsion for cars in its infancy, the final mixing and matching of the various electric power options will depend as much on new and refined technological developments as it will on consumer demands. However, it is being increasingly recognized that the PEFC will become the system of choice for the primary power source. Finally, as the raft of economic, environmental and political issues associated with the burgeoning use of increasingly limited liquid petroleum reserves unfurl over the coming decades, the move to electric propulsion systems for cars will inevitably hasten the development of new and improved PEFC and associated electric power systems.

1. Kerr, R. A., *Science*, 1998, **281,** 1128.
2. Campbell, C. J. and Laherrère, J. H., *Sci. Am.*, March 1998, 60–65.
3. Hatfield, C. B., *Nature*, 1997, **387,** 121.
4. Hiller, K., *Erdöl Erdgas Kohle*, 1997, **113,** 349.
5. *Electric Vehicles: Technology, Performance and Potential*, OECD/ IEA Publication, 1993.
6. Gosden, D. F., *Trans. IEAust.*, 1991, **GE15,** 83.
7. Hoogers, G., *Phys. World*, August 1998, 31–38.

Received 29 June 1999; accepted 27 August 1999

The dharma of ecology

T. N. Khoshoo

The concept and scope of ecology is ever-widening and becoming all encompassing. The ecological crises facing the world are basically an outward expression of inner crises in the mind and the spirit of human race. This species has changed landscapes on earth beyond recognition for its own 'good'. There is now a realization about interdependence between welfare ecology and welfare economics. Both have to be fortified by ecological and technological assets. Ecology is becoming a moral issue and has a deep interconnection with dharma. Proper interface among ecology, economy and technology will lead to welfare of biosphere of which human being is an integral part along with all other living creatures and non-living materials.

Ecology: A moral issue

The word *Dharma* enjoys universal acceptance having been included in all the standard English dictionaries. Now it is as much an English word as it is a Sanskrit word. It is derived from the Sanskrit root *dri* which means to 'uphold, sustain and support'. In simple language it means to 'hold together the different aspects and qualities of a being'. Associated with it are also righteousness, morality and duty. In short, it embodies all that is universally and eternally true. Without dharma nothing can make sense. Therefore, it is a part of the very nature of every thinking human being about all situations and problems (including ecology) that confront humanity at large. Dharma is, therefore, enshrined in any orderly life, society and environment. Implicit in it is that human beings have to control themselves so that their actions do not endanger the ecology which surrounds them, and on which they depend for sustenance all the time. Also implicit in dharma is that one should not inflict on surroundings and other living beings anything which is disagreeable to one's own self. Thus, there is a deep inter-connectedness between dharma, ecology and environment that surround all forms of life all the time. In view of this, it is not surprising that ecology and environment are fast becoming moral issues and a moral responsibility of the human race which has the capability to think and foresee about the end-result of human actions. Nature (*Prakrati*) and human being (*Purush*) are two major elements recognized in the scriptures, which, if antagonistic, can bring doom and gloom to the Mother Earth.

Normally nature by itself does not degrade environment. If, however, natural cataclysmic changes happen, there may follow environmental degradation. Left to Nature, there starts a process of ecological rehabilitation and reconstruction of the deteriorated habitats, and, more often, a new ecological regime sets in, which may even bring *status quo ante* in course of time, or even a new balanced ecological state. There is, therefore, tremendous resilience in Nature, because of the inherent capacity to reconstruct and rehabilitate. Nature is also not static, because there is an inherent capacity in it to change, refine and update. Those of us who visit natural habitats see these phenomena occurring all the time.

On the other hand, market forces, more often than not, depend on short-term gains and profits. These are oblivious of the responsibility of setting right the damage created by their short-sighted policies. Regrettably, at present making profits is the dharma of industry, but losses regarding generation of wastes/pollution is governmental and societal responsibility. Even at the individual level, eating food every day is a personal matter, but disposal of wastes therefrom is societal and/or governmental problem. Environment is the source of all raw materials which everyone is out to grab, but environment is also the sink for all wastes. A question arises as to how moral are such attitudes? Therefore, benefits and costs must become part of all environmentalism.

The world is not united on the question of sustainability of the Earth system including a concern about growing human numbers. However, most scientists are worried about the shape of things to come. They advocate understanding the basic questions scientifically and evolve technologies to combat the impeding dangers. Earth being a finite entity, does not grow in size. Thus there is a need to combat realistically the problem of increasing human numbers, and their wants and desires, and qualitative and quantitative dwindling of resources and above all the very health of the Earth system.

The basic question is, can we raise the carrying capacity of the Earth system to cope with demands of *one* species (*Homo sapiens*) which happens to be the pinnacle of organic evolution! Using technological innovations, this species having spoiled the Mother Earth, no doubt has also the technical capability to stop endangering the health of the Earth system! This sounds paradoxical, but is nevertheless true.

The above are some inconvenient but real-life questions for which we have to find answers: sooner the better. Here then is a combined challenge for scientists, technologists, economists, sociologists, and those who deal with ethics of resource use. The basic question arises as to what will confer sustainability! Some thinkers (including this author) have attempted to answer this, but

T. N. Khoshoo is in Tata Energy Research Institute, Darbari Seth Block, Habitat Place, Lodhi Road, New Delhi 110 003, India.
e-mail: khoshoo@teri.res.in

there is need for a more concerted attention of an expert group so that a necessary policy frame-work can be drawn for this purpose.

To save our planet with all its living and non-living manifestations and to ensure the diversity that has been its strength, there is an urgent need to adopt a Code, which may be called *The Dharma of Ecology*. Without following dharma nothing can make sense. Human being is a thinking species, therefore, dharma has to be part of its very nature including the ecology that surrounds it. Although this word is an oriental coinage, it is universal in approach and application. It is connected with human conduct and is enshrined in all religions of the world in one form or another (Khoshoo, unpublished). The important point is that all living and non-living resources in the life-support system are held in an intricate balance and have a value. These resources are to be held in trust. Thus human action should not inflict on other species (including other human beings) anything that is disagreeable to one's ownself including the surroundings of a particular individual be it plant, animal or micro-organism.

Some basic principles

The following are some basic principles underlying the dharma of ecology:

- Protecting and augmenting the regenerability of life-support system. This has to be accomplished by rationalized husbanding of all resources. Among other things, this would involve nurturing and protecting renewable resources; conserving non-renewable ones together with prolonging their life by recycling and reuse; avoiding waste; and benefiting from the economy of scale.
- Fair sharing of the resources, and means and products of development between and within nations of the world. This would reduce the disparity in resource-use, leading to a significant reduction in resource-use in the developed countries and increase in resource-use with little or no environmental degradation in the developing countries.
- Promoting awareness regarding the hidden social, economic and environmental costs of consumerism and overuse of resources with particular reference to its impact on the developing countries.
- Adopting willingly sustainability as a way of life by encouraging *frugality*, i.e. getting more from less, and *fraternity*, i.e. getting it in association with others.
- Meeting genuine societal needs and legitimate aspirations of the people by blending economic and environmental imperatives so as to alleviate poverty.
- Halting and then reversing the overuse of resources and armament build-up for ensuring sustainable environment, peace and security.

We need a firm commitment to the dharma of ecology at the individual level, because a society or a government is only an extension of an individual. The common threats to the long-range ecological security will bring nations of the world together. The Earth as a whole is also a Civilization Reserve not only for humankind but also for all the living beings: be it plants, animals or micro-organisms. Therefore, as citizens of the world, the human race must rise above the local and national ideologies and narrow economic systems, and owe allegiance to the life-support-system as a whole.

Global family

Never before, has there been a greater need for application of the concept of Global Family (*Vasudaivakutumbakam*) as is today. Environmental crises facing the world are actually an outward manifestation of an inner crisis in mind and spirit of human beings. Environment can no longer be treated as bits and pieces and dealing only with wildlife, ecodegradation, pollution and the likes of these. In the larger context, environment encompasses the whole well-being of all life on our planet. In the developing countries, poverty is the biggest polluter, a statement made by the late Indira Gandhi. Poverty degrades environment and thereby accelerates the pace of poverty in the developing countries. Their dire need is a survival strategy. On the contrary, in the developed world, it is the prosperity and unlimited greed which causes environmental degradation. Even though the developing countries harbour over 84% of the people, their contribution to ecodegradation and pollution is far less than that by 16% of the people in the developed countries, who consume nearly 80% of the world's resources.

If history of human being is traced ever since its origin in Africa, it is clear that, from the environmental and socio-economic points of view, there were three major societal epochs discernable: the Hunter–Gatherer Societies, followed by Agricultural Societies and the more recent Industrial Societies. We may now examine the broad contours of each of these.

Hunter and gatherer societies

The human being has been a hunter–gatherer for 99% of its time span. It is only during the last ten to twelve thousand years that it has taken to agriculturization and industrialization. During the hunting–gathering stage, the human being was largely nomadic, and acted as one of the species in the concerned ecosystems. The environmental impact was strictly local and small, and due to the natural process of eco-repair, ultimately there was little or no damage. Hunter–gatherers have performed the biggest trial-and-error experiments for the humanity as a whole. The latter has to be ever-grateful to the former.

Agricultural societies

The early agricultural societies domesticated livestock for food, clothing and for carrying loads. They also began selecting and cultivating plants as food in 12 centers of origin and domestication in the world, one of which is in India. Except for some micro-organisms, humankind has not added to the list, and has been using the same animals and plants that were selected and domesticated by its primitive ancestors. However, with the invention of the plough and the wheel, agricultural societies were involved increasingly in clearing forests for cultivating crops, raising livestock and making dwellings. With rather assured food supply, population began to increase and food supply had to keep pace with it. Thus irrigation helped in settled and enhanced agriculture in turn leading to significant increase in population and permanent settlements in the form of villages, some of which in the course of time became towns. Some of the towns grew into cities.

Together all these factors resulted in the establishment of civilizations. Associated with the latter was enhanced need for food, leading to enhanced rate of degradation of forest cover, and considerable increase in irrigation systems. The latter began to become clogged due to siltation and associated environment and human health problems followed.

Since, by now, population had begun to increase and agriculture had extended considerably, there was need for labour both for agricultural and desilting operations. This gave birth to a landed class who owned land, and a landless slave or labour class who put in hard work. The small and localized environmental impacts gave way to larger impacts on account of forest clearances for agricultural purposes and grasslands for domesticated cattle. The human being still depended on its muscle power and that of the domesticated animals.

Next came the *Agriculture-based Urban Societies*, which led to further increase in population. Moreover, while some villages produced food, the larger villages grew into towns and larger towns into cities. In the latter case, people depended on food produced in the villages. In their spare time in cities, people took to small industries like tool-making, weaving, pottery, handmade goods, etc. Six such contemporary civilizations appeared on the Earth; these were Nile Valley, Babylonian, Greek & Roman, Indus Valley, Huang Ho Valley, and Mayan & Aztec. While these civilizations contributed materially to literature, art, music, science, etc., there were two classes in each: the *haves*, who constituted a small section but had large assets and were powerful; and the *have nots*, who were a large section, with little or no assets and were powerless being involved in producing food and doing all the dirty work and rendering services all the time.

Earlier, fights between groups took place for possession of more and more livestock, but now fights began about the ownership of land. This led to the springing up of leaders with armies of followers who controlled large areas. Wars began to be fought for possession and control of land and ecological assets. There was scant respect for assets like water, forests and land which were poorly managed and overgrazed, resulting in soil erosion, blockage of irrigation systems and increased number of slaves to clear the silt. The cities had a lot of waste generated by people, leading to infectious diseases and parasitic attacks. Habitats began to be altered beyond their carrying capacity, and, for the first time, there was significant ecodegradation. In this process, some empires became weak and wars became frequent. All this resulted in further degradation of the environment. Such ecological, economic and social reasons led to the collapse of the six civilizations enumerated earlier. In short, the prime reason for the collapse of civilizations has been disrespect for forests in particular and environment in general.

Industrial societies

Starting from England, in the Western Europe was born the Industrial Revolution, with many inventions involving coal-based steam engine systems followed by the internal combustion engines. Thereafter, horse carriages and wind-powered ships were replaced by engines using fossil fuels. This was the period of European expansionism into Asia, Africa and the Americas. In this process, the indigenous peoples were either largely annihilated or subjugated.

Even agriculture now began to be based on coal and oil in place of human and animal energy. Production increased and there was migration of former farmers to towns and cities. They now took jobs in mechanized factories. With the two world wars, fought in the 20th century, many inventions were made in the area of science and technology. After the wars, these led to mass production of useful products at affordable prices and a ‘high’ standard of living with higher GNP per capita. With the application of modern science and technology, there have been major gains in the yield potential of the domesticates. There also was improved life expectancy, better living conditions, education and old age security.

The environmental impacts of the industrialized societies were tremendous, be it agriculture, industry, mining, etc. All these led to degradation of land, forests, water, biodiversity and air through the release of noxious chemicals and cutting down of forests. Most cities became twin cities, the mega-component with all the facilities, and the slum-component where ecological refugees live. Most cities in the world are still stuck with such a situation. There also developed the regional problems of acidification and global build-up of carbon dioxide and depletion of ozone.

In fact, industrialization has been a mixed blessing. There was considerable economic growth with per capita increase in GNP and overall standard of living. However, all this progress and benefits have been at tremendous environmental costs. Furthermore, for some time past, lifestyles in the developed countries have also affected the resource base in the developing countries. The classical cases are that in return for food and financial aid by the developed to the developing countries, the latter destroyed their forests by supplying timber, growing cash crops and producing cheap meat for consumption in the developed world. In this regard the well-known case is the *Hamburger Connection* where Norman Myers showed that 40% of the forest cover in Central America had been destroyed for making pasture land available so as to supply beef at cheap rates to North America. The present-day cost of beef does not reflect the *true cost* of its production because huge environmental costs are not added to it. This example stirred the conscience of the whole world. The developing countries also use obsolete and dirty technology supplied by the developed countries, thus degrading the environment further. In return for financial aid, some developing countries have even offered sites for burying and dumping noxious wastes. All such aids are in fact *concealed compulsions* and, in practice, amount to acts that threaten the ecological security of the poor developing countries.

Thus, in the developed countries the causes of eco-degradation and pollution are their prosperity and greed, while in the developing countries the causes are poverty and need. In the latter case, it is matter of very survival. The most profound aspect of the industrial era has been the arrogance of humankind to consider itself the most superior organism in the biosphere, and a growing feeling that everything is subordinate to human needs, and a feeling of being a co-creator.

Today the world is rather divided into two camps: a few (26) developed countries mostly located in the temperate regions of the world and a large number (107) of developing ones in the tropical, subtropical and hot temperate belt. The former consume far more resources (over 80%) than the latter. The underlying feeling of undue exploitation of resources by the developed countries exists in the developing ones. This causes tension and friction. However, in the recent years, the developed countries, confined mostly to the temperate regions, have realized the criticality of tropics and subtropics for their own survival and well-being. This has led to a trend to swap the debts of the countries in the tropics, for conservation of tropical forests. It is indeed a healthy sign, because environmental interconnectedness and interdependence between the rich and the poor nations is becoming increasingly clear. No nation however rich or poor is safe if its environment deteriorates significantly.

Environmental problems are thus the result of interaction between complex and poorly understood social, economic, technological and political factors. However, it is also clear that although developing countries suffer from problems of over population and lack of resources, the net quantum of eco-degradation and pollution in their case is far less than the less-populated developed countries. Furthermore, pollution in the developing countries is mostly biodegradable, while that in the developed countries is mostly non-degradable.

Ecological ethics

In the coming years it is certain that ecological ethics will get added importance. The Western religions (Judeo-Christianity, Islam and Zoroastrianism) have by and large looked at the relationship between humankind and Nature with a measure of arrogance and an underlying co-creator attitude: A notable exception being St. Francis of Assisi. The result has been *conflict* with Nature. On the other hand, the Eastern religions (Hinduism, Buddhism, Jainism, Sikhism and Taoism) have overwhelmingly viewed environment and Nature with reverence and an underlying partnership, leading to *harmony* with Nature. Most orientalists start their day with prayers to Nature and the bounties it offers. The two components Nature (*Prakrati*) and humankind (*Purush*) are partners which must work harmoniously.

A lot of useful literature is now emanating from the western world about the ethics of resource use because, more than the east, the west has realized that their present-day pattern of development is not sustainable. They are eager to hear the views of orientalists about the environment, because this subject has been a part of ethos of the latter from time immemorial.

Connected with the subject of ecological ethics is the fact that the human race has had a common origin (in East Africa) and also a common past. Then there followed divergence, and human being colonized all the continents because it was the first intelligent, inquisitive and thinking animal (Figure 1). In due course of time, there followed population explosion, multiplication of needs, undue demands on and progressive destruction of components of the Earth system (namely: atmo-, hydro-, litho-, and biosphere including biodiversity). The net result has been that the Earth system as a whole became progressively endangered: some of its parts more than the other parts.

Then there began a global realization about the impeding dangers associated with serious environmental deterioration. Then came the Stockholm Conference (1972), followed by the Rio Conference (1992), and a plethora of other conferences. In this process, humanity as a whole jumped from *Common Origin* to the concept of *Common Future* (Figure 1). There has been talk of globality of environment, and connectivity between local and global environments. Yet there are no worthwhile global or regional strategies or even national strategies for achieving sustainability. Therefore, while *Common Origin is a* fact, *Common Future* is still a myth (Figure 1). Some years ago, M. S. Swaminathan raised a very pertinent question: How can there be a common future without a common present? The latter is still an open question and an enigma! Should not humanity do something tangible about it? This is a moot question which needs to be addressed to very seriously.

The only option left to the human race is to not only work out solutions to local problems, but also to rise above the local issues and think about the repercussions of these at the national, regional and global levels. Furthermore, it has to work over-time to give all such strategies a practical shape. It is indeed a two-way traffic. Understanding the dynamics of this two-way traffic will actually lead us towards real sustainability in development.

Apostles of ecological dharma

Regrettably during the 20th century, the human race has seen more tormenters (at least four) but only one benefactor (Mahatma Gandhi). In recent times, three Indians who, in every sense, preached and practised the *Dharma of Ecology* are: Mahatma Gandhi, Vinoba Bhave and Mother Teresa. The former two were Indians by birth but the last one was by her voluntary adoption. In fact all the three belonged to the whole humanity. The first two were devout Hindus, the last a devout Christian. But all the three followed identical paths and reached similar conclusions: to care for the poor, the dispossessed, the deprived and the destitute or, as M. S. Swaminathan has said in a different context: *reaching the hitherto unreached*. Thus, it was sheer simplicity that these three great souls wore. Here then are ideals in sustainability for the whole humanity.

The lessons one draws from the past experience are loud and clear and there is considerable realization about the following:

- Earth is a finite system, both in resources and in its carrying capacity;
- Future economic growth cannot be sustainable if it is at the expense of long-range ecological security;
- Environmental insecurity ultimately leads to economic, social and political insecurity;
- Sustainable development for intra- and intergenerational human well-being has now to be an integral part of the future composite world culture; and
- Sustainability in development is a global concept and every living being, as a member of the World Family (*Vasudaivakutumbakam*), has a role to play.

There is an urgent need to translate these lessons into reality through the *Dharma of Ecology*. While we must understand scientific and technical complexities of nature, we must not do so with arrogance of *conquering* nature, but *working in close harmony* with it. We must develop a good measure of reverence for nature for the vast bounties it provides. In this connection, we must also learn from the tribal societies, which have developed an approach of harmony with nature. This can still be seen in the interiors of the Andaman and Nicobar Islands and Amazonian forests.

If there is any one thing that is going to bring nations of the world together, it is the *common threat* to our long-range ecological security. Therefore, before we talk of common future, there is need for common concerns, approaches, strategies and actions for our *common present*. Thus, for our sustainable future, we have to move towards globality on the one hand so as to correct the environmental follies, particularly of the industrial countries; and on the other hand, we need to meet common global threats. There is need to develop a culture/ ethics/code for Ecological Dharma at all levels starting from the individual up to a country or region and the entire globe so as to practice the cult of sustainability in development. It is only then that we will have a situation as put by Rene Dubos: '*think globally but act locally*'.

A basic question arises: Are we moving towards a sustainable society? This indeed is a major challenge as also an opportunity before the entire human race. In India, if we go on the way we have been so far, on 1 January 2001 like today, centuries will continue to co-exist. We will continue to have a subsistence India of a large number of poor and dispossessed toilers and plodders who live in medieval times, and an affluent India of a small number of people who are jet-set and wealthy. The latter may be poised to enter the 21st century with a bang. How soon we take even the preliminary steps to bridge the vast gap between the large but powerless subsistence and the small but powerful affluent India, will actually determine whether we can make it to a sustainable society, where we have environmental harmony, economic efficiency, resource conservation, gender equality, equity with social justice, and local self-reliance. To practice this, we need to draw inspiration from Mahatma Gandhi, Vinoba Bhave, and Mother Teresa.

Future prospects: Welfare ecology

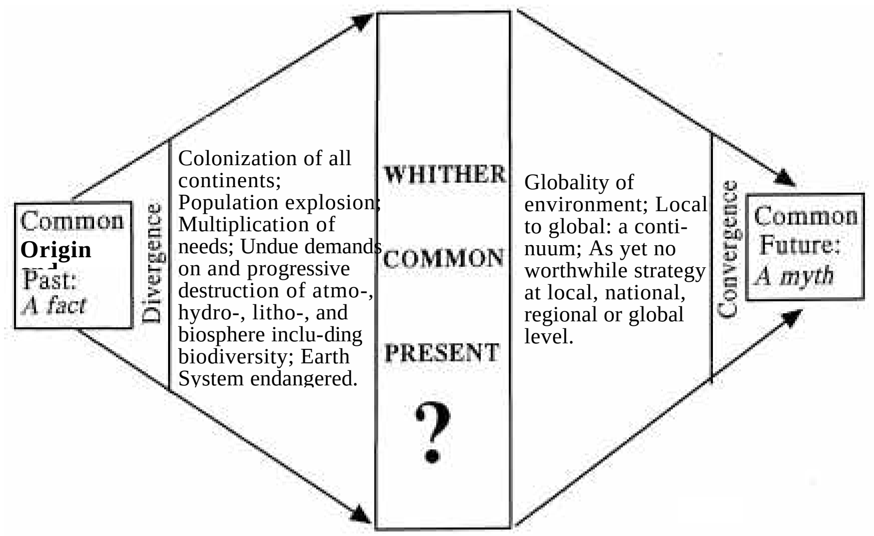

Figure 1. Transition from common origin to common future.

Thanks to Dhrubajyoti Ghosh an all-encompasing term, *welfare ecology*, has now been introduced in ecological literature (*Selected Essays on Welfare Ecology*, Centre for Sustainable Living, Calcutta). This is a sequel to Amartya Sen's welfare economics, who, for the first time, talked of economics of the weak, the dispossessed, the deprived and the destitute which constitute the dumb majority in any developing country. The strength of a chain is its weakest link, and, therefore the poorer section in any society must receive special attention. Once the teaming millions come out of the morass of poverty, penury, illiteracy, hunger and dire want, then only a developing country can progress as a whole. Therefore, welfare economics has to be backed by welfare ecology. A basic premise is that economy springs from the use of ecological assets (atmosphere, hydrosphere, lithosphere and biosphere) coupled with human ingenuity in the form of technology (Figure 2). It may, however, be pointed out that technology is not only a human attribute but many other organisms make use of it intuitively. For instance, one has only to have a mind and an eye to see how meticulously and efficiently bees are organized socially and build their hives, how birds build nests, or how a beaver (an amphibious broad-tailed soft-furred rodent) builds a dam in a gushing stream of cold water. A bee-hive is an example of one of the most perfect and articulated organization. Each bee knows its job which it does selflessly. These are marvels of technology, division of labour and perfect coordination and articulation, in no way less than human ingenuity, if not better because there is no element of personal greed. Thus welfare economy and welfare ecology are mutually supportive. Gone is the time when ecology meant only study of plants and animals in their habitats, more often such discourses included human being very marginally.

Human ecology is now an important subject. There is a deep interconnection between human needs, wants and aspirations which in the wealthier sections of any society are in reality unlimited. Therefore, there is an urgent need for the human race to address itself to a serious question like: what is *enough* for a simple but comfortable lifestyle avoiding ostentatious and vulgar show of wealth which causes undue stress on environment and waste of materials? Welfare ecology is relevant to all living organisms including human being. It embraces the whole biota, because the health of whole will determine the health of the part, and vice-versa. Therefore, welfare ecology has a very wide meaning and application. Inherent in it is the basic minimal requirement for a simple and comfortable lifestyle which can be permanent with no long- or short-range ill-effects on the environment in which an organism lives. Sustainability will become a reality only when one lives on the mean annual increment (MAI) of the basic ecological–economic capital.

Thus there is a deep interconnection, interdependence and interrelatedness between welfare economy and welfare ecology. The two are mutually supportive. On such a mutuality depends the future of humankind on a sustainable basis. Proper interface between ecology, economy and technology, will lead to welfare of biosphere of which human being is an integral part along with all other living creatures and non-living materials. We need to face ecological challenges of the 21st century with the joint message of welfare ecology backed by welfare economics and vice versa.

Economics, energy and ecology are also interrelated, and one of the major causes behind India's environmental problems can be traced to their bad management. At present only economics plays an overriding role even when ecology is actually regarded as biological economics and energy as a currency of life. As of now three major questions confront humanity. These are: How can the huge ecological deficit already with us be wiped out without adding to the present-day ecological problems? How can the future development be made sustainable? How can aims and objectives of environment and economic development be reconciled and be unified?

Conclusions

Although a Sanskrit word, dharma is now universally accepted, it has a deep interface with ecology. Among the important findings of this century is the fact that the Earth is the only planet in our planetary system that supports life as we know. It is our only home. All the living beings (plants, animals and micro-organisms) on Mother Earth constitute one Global Family. Furthermore, the 20th century has been one of discovery and expansion of human activities, resulting also in considerable environmental destruction. On account of this, the human race by its action has been responsible for extinction of some of the life forms. A question arises whether the next century will be one of continued and rapid environmental destruction, or of environmental reconstruction so as to save as many life forms as possible and the Planet Earth as a whole? Humankind has to make up its mind about becoming more humane and less selfish. There is an urgent need to ensure continued regenerability of the life support systems, to be

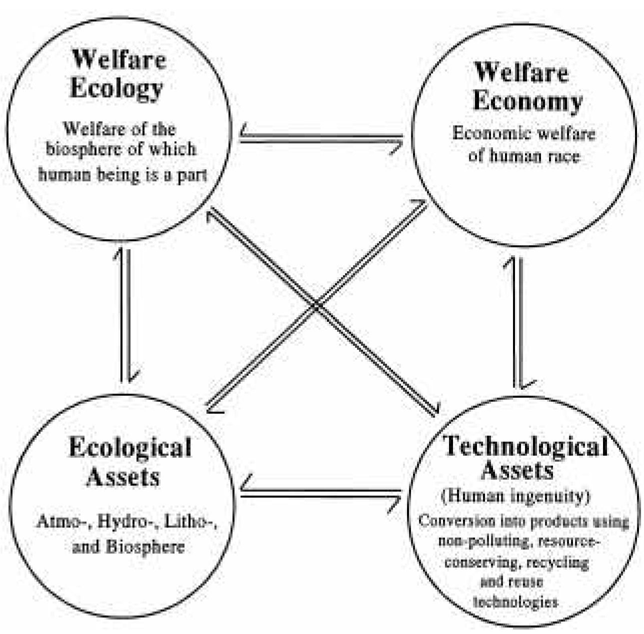

Figure 2. Interrelationship between welfare economics, welfare ecology, and ecological and technological assets.

followed by fair-sharing of resources and their products, and practising frugality, fraternity and sustainability. Adopting such a course of action would help answering a basic question: how much is *enough* for a simple need-based comfortable lifestyle? In turn it would also help stall the ecological decline that has already set in, which if unheeded would in turn lead to economic decline followed by social disintegration. History has been a witness to such a course of events. Before any civil society talks about common future, it has to ensure a sustainable present. To attain the latter would need inputs from all sciences, technology, socio-economics, ethics and law. There is, therefore, a need for an indepth thinking on these issues.

We need to draw lessons from the decline of once flourishing civilizations in the medieval times, and avoid disrespect for Nature at all costs. We also need to conserve not only the natural heritage, but also the intellectual heritage. In the natural heritage is included the Mother Earth itself with all the biomes, ecosystems and populations of all living species (including the human being). In the intellectual heritage is included all that has been crafted and created by human genius for the good, the benefit and the well-being of humanity at large. It would also include human settlements, science and technology, history, culture, religion, philosophy, art, literature, music and dance, handicrafts, myths, etc.

The civil society needs to be commited to make innovations in development possible and thus ensure a better life for the generations to follow and help in sharing and caring. Herein lies a dual responsibility for each one of us: one to the biosphere and the other to humanity and all life forms on a collective basis. In short, there is need to guarantee a healthy Earth by itself, and the life on Earth in all its manifestations.

To conclude, sustainability is not only a scientific, technological, social, and economic issue, it also has major moral and ethical dimensions. Welfare economics backed by welfare ecology together hold the key to human survival on a sustainable basis. Therefore, determined efforts have to be made to avoid crossing the thin line dividing sustainability and unsustainability. To achieve this, there is also a need for evolving a unique ‘technology’ for the ‘inner’ development of human kind itself so that misuse of resources and creation of unsustainability is avoided. To the present author, these are some of the basic and *dharmic* responsibilities of humanity as a whole.

ACKNOWLEDGEMENT. I thank Dr R. K. Pachauri for providing facilities.

Received 8 July 1999; revised accepted 25 August 1999

Conservation, management and documentation of livestock genetic resources and biodiversity: A biotechnology perspective

*Sher Ali and Seyed Ehtesham Hasnain**

For conservation and management of livestock genetic resources, in-depth molecular characterization of a species is required. This may encompass genome organization, sex-determination, regulation of gene expression, clad identification, breed delineation, assessment of QTL loci and their synteny in related species. Molecular methods are available that may be employed for genome analysis of farm animals as well as highly endangered species. We discuss various methodological approaches together with molecular basis of the development of genetic markers and marker systems in the context of genetic conservation.

THE conservation and management of India's immense genetic resources require great deal of effort and sustained multipronged approaches. Extensive survey, cataloguing, physical evaluation, in-depth characterization, preservation, maintenance, timely restoration and sustainable utilization of livestock genetic resources without impinging upon their delicate ecological balances and inter-clad interactions are important in addressing these issues. In the entire spectrum of cognate activity, in-depth characterization will involve many scientific approaches for which genetics and biotechnology will continue to play dominant roles.

The gigantic spectrum of biodiversity and the fast-changing world economic scenario coupled with politically ever-unfolding newer global order makes it imperative to take a serious view of our current bio-assets and unforeseen future liabilities. A marked number of breeds of cattle, buffalo, sheep, goat, pig, horse, camel, mithun, yak, dog, cat, poultry duck, geese, turkey, guinea fowl, pheasant and many more have been reported but most of them remain genetically uncharacterized. A battery of molecular markers and different experimental approaches are required not only for the genetic characterization of the species that are related to our economy but also to those that are part of our eco-system[1–4].

Startling as it may appear, but the fact remains that commensurate to the abundance of biodiversity, genetical data at every level is lacking. Much of genetical or biological information may be available on experimental animals such as rat, mouse and *Drosophila* but the same with respect to farm animals such as cattle, goat, sheep or mithun are rare. While this does not mean that studies on experimental animals in any way are less important, it does reflect lack of effort towards the understanding of the genetics of farm animals and other species that are part of critical mass of our biodiversity. The generation of detailed genetical information in turn would involve genome analysis with respect to sex determination, gene expression and its regulation, and the genomic organization and expression of the sequences involved in signal transduction. Similarly, studies on protooncogenes and their detailed characterization, molecular events leading to controlled cell proliferation and apoptosis would prove to be useful for undertaking advance studies on the desired species. Keeping in view the physical and phenotypic attributes, clad identification and their proper phylogenetic positioning[5,6], breed

Sher Ali and Seyed Ehtesham Hasnain are in the National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India; Seyed Ehtesham Hasnain is also in the Centre for DNA Fingerprinting and Diagnostics, CCMB Campus, Uppal Road, Hyderabad 500 007, India.

*For correspondence. (e-mail: seyedhasnain@hotmail.com)

delineation[7], assessment of QTL loci and their synteny in related species, mutational studies[8–10] and comparative analysis of molecular events leading to modulation of signal transduction are other much-needed areas of foci in the animal species. Similarly, sex identification, analysis of the genes and molecular events involved in sex determination will be very important. Information on these lines would provide the base for subsequent manipulation of genomes related to transgenic experiments, even 'tailor-made' animals and eventually for the propagation of desired germplasm.

In order to achieve the above-mentioned goals, several experimental approaches can be adopted keeping in view the fact that a typical higher vertebrate genome is endowed with 3–4 billion haploid sequences representing at least 80,000 functional genes including a sizable part of non-coding repetitive DNA. Identification of gene(s) associated with genetic superiority of the species may be a challenging task but the same is not insurmountable owing to the availability of many powerful techniques of molecular biology.

Recent studies have shown that a number of repetitive DNA belonging to different families are transcriptionally active, thereby, refuting the earlier contention that repetitive DNA is 'junk'. Repetitive DNA may in fact prove to be useful markers or may even assist in the development of marker systems[4,11–16]. The fact that an individual chromosome is a uninemic entity (made up of a single DNA molecule running all along the length of the chromosome), it becomes important to understand the organizational and expressional relationship of a functional gene in the context of organization of repetitive DNA. It is however, not clear if the intergenic distances with respect to different types of satellite DNA are always constant and that the similar types of spacer sequences are involved in the intergenic organization, at least in related species, if not in all the species. Information on the presence of different possible types of satellites in a species, their possible association/linkage within and between other related species will be desirable. Finally, a comprehensive synteny amongst all the desired species may be established. Such endeavour would in the long run provide working flexibilities to use marker(s) originating from one species in another. This may also help to provide an insight into the genetic susceptibility or possible resistance of a species to a disease.

In this context, the following technological approaches, though not necessarily in the same order, may be adopted and molecular genetical studies be directed towards some important species such as cow, buffalo, goat and sheep which can be subsequently, extended to other species.

To establish and ascertain different breeds of animals

A number of breeds based on physical attributes have been reported. This aspect may be substantiated scientifically for which a variety of standard experimental protocols are available. Monolocus and multilocus DNA profiling[1,3,7,17–22], random amplification of polymorphic DNA (RAPD), also known as arbitrarily primed polymerase chain reaction (AP-PCR) or more robust approach like minisatellite associated sequence amplification (MASA) may be used[23–25]. For the primers needed for RAPD or MASA, individual hypervariable repeat loci may be cloned, sequenced and screened for the presence of short tandem repeat motifs (STRs) endowed therein. These STRs in the range of 6–18 base residues may be used for RAPD or MASA studies. Some of the clones generated in the process may uncover high level of genetic polymorphism and may prove to be equally useful for routine DNA fingerprinting. A STR-based probe useful for DNA fingerprinting will also be able to score all the associated beneficial points such as pedigree analysis[5,15,26–28], allele drop out, assessing rate of mutation[8,9], studying population structure, monitoring successful breeding partners, ascertaining the propagation of desired germplasm, assessing the gene pool for its overall heterogeneity[2–4,29] and dilution of the same owing to infusion of sudden newer genetic materials in a dynamic population[2,3]. This would also enable to uncover the effect of genetic drift and migration, since allele length frequencies are likely to be affected. Similarly, some of the clones so generated may prove to be useful for breed delineation following simple DNA typing system involving Southern blot hybridization. Different recombinant clones originating from a single satellite fraction comprising consensus sequences of about 15–30 base pair may be of particular biological significance. Oligonucleotides based on these consensus sequences may be used directly as probes to uncover species or breed-specific pattern following dot blot or Southern blot hybridization. Similarly, complete sequencing of about 15–20 independent recombinant clones representing a particular satellite fraction followed by their alignment would uncover polar or non-polar mutations. Analysis of such mutations may also form the basis of breed delineation and may even be linked for their uniqueness with respect to a population. Usually, such non-polar mutations are extremely polymorphic and may reveal micro-level changes in the population[5]. This would also provide a model system to study the environmental effects on the animal population originating from two diverse environments[30,31].

Using the above-generated recombinant clone(s), expression studies may be conducted with the total RNA isolated from different somatic and germline tissues within and between individual animals employing northern-blot analysis. Subsequently, depending upon the detection of signals, the data may be substantiated by RT-PCR experiments as detailed in the following paragraph.

To explore possible transcriptional status of the satellite DNA or satellite-linked sequences

Sizable body of literature is available on different types of satellite fractions from different species including human. However, a systematic search towards the association of these sequences with functional gene or towards their own transcriptional status has never been attempted in the farm animals. This is corroborated by the fact that while a total of 1745 homologous loci in humans and 240 in cattle have been reported, only two have been characterized in the bubaline genome.

An extensive enzymatic restriction survey of the DNA (of the desired species) and identification of satellite fraction(s), their cloning and sequencing will be the first step. The sequencing of the contig will allow primers to be designed for other studies such as monitoring copy number variation, their expression by Northern analysis and RT-PCR-based expression in both the sexes during the course of development in different somatic and germline tissues. It is envisaged that a number of satellite fractions may be found to be transcribing. Analysis of these sequences using *Blast Search* and their homology with entries in the GenBank would help identifying if the cloned satellite were associated with a known functional gene or the same represents still uncharacterized coding sequences. Exclusion studies using these clones may help narrowing the search of much-desired QTL loci that have *hitherto* been an elusive proposition for most of the physical and physiological attributes. Molecular characterization of a large number of satellites following the above-mentioned approaches will maximize the chances of saturating the genome with appropriate markers and marker systems.

Studies on genes and repetitive sequences implicated with sex determination and mutations known to cause infertility associated with heteromorphic sex chromosome(s)

These studies have three apparently unrelated but actually linked components. A typical vertebrate species usually has heteromorphic sex chromosomes with sizable portion of heterochromatin comprising various types of repetitive DNA. In the context of animal biotechnology, for sex identification, use of a marker that uncovers differential organization of the heterochromatic sequences (even without its involvement in actual sex determination) is possible[32,33]. Earlier work has shown that several types of repetitive DNA are involved in the phenomenon of heterochromatinization. However, it is not known if the heterochromatin in all the species is qualitatively and quantitatively similar and the same is linked with the evolutionary status of the species as has been suggested by some workers. Sex identification in a species can be conducted either by using: (i) repetitive DNA probe for dot blot or southern blot hybridization or (ii) primer(s) that amplify the sequences exclusively from the heteromorphic sex chromosome or (iii) primer(s) that amplify sex specific haploid gene(s). The genes responsible for sex determination and associated sequences differ in different species. This warrants an in-depth analysis of this phenomenon in different animal systems providing ample scope for undertaking research on these lines. The envisaged outcome would eventually assist in deciphering fertility status of animals and identifying most prevalent causative factors leading to infertility in a given population. Sex identification in itself may be a small exercise conducted routinely, but it does not tell the actual fertility status of a species. Thus, a detailed genetical study needs to be carried out to be able to establish fertility status involving all the possible genes. The rationale is based on the fact that mutation(s) known to affect fertility status of a species may involve different genes. Once most commonly occurring mutations become known, the same may be studied in a given animal population which may then be used to ascertain if the fertility/infertility situation is similar in different animal populations with respect to particular loci. Deviation, if any, may be attributed to non-genetic factors and this will open up newer vistas for additional research.

Studies on the autosomal gene(s) in the context of animal fertility

It is well established that in many vertebrates, genes located on the sex chromosome are neither sufficient nor responsible for sex determination. A number of autosomal genes have been implicated in the cascade of events leading to organic sex determination. One such autosomal pleiotropic protooncogene *c-kit* receptor has been reported to be implicated in haematopoiesis, melanogenesis and gametogenesis. Multiple mRNA transcripts of this gene and alternate splicing have been reported[34]. In recent studies, intron/exon reshuffling has been observed. In rat model system, the mRNA transcripts analysed from normal and infertile ones have shown one amino acid deletion in testes mRNA transcript compared to somatic tissues including brain. This observation is substantiated by histological studies on testes. Identification of mRNA transcript involved in gametogenesis in normal animal and its mutant forms leading to infertility would be of practical implications for the farm animals. Similarly, studies on other genes known to be involved in sex determination such as SRY, ZFY or Sox family genes would be fruitful. A collective approach towards the understanding of the phenomenon of sex determination and expression of involved genes and their mutant alleles in a given animal population will help towards conservation of biodiversity and management of genetic resources. Further, this information may be useful for reshuffling the two populations to upgrade the relatively confined gene pool of genetically less heterogeneous populations[2–4].

Studies on the rare genetic resources

There are a number of animal species endemic to India such as pigmy hog, dwarf cattle, or mithun which require genetic as well as population-based study in the light of the ever-growing menace of bio-piracy. While local research laboratories may not be fully equipped to undertake this daunting task, effective collaborative research project(s) may be undertaken involving the better equipped (nodal) and regional laboratories. This, while generating the much-needed vital information will also augment the regional infrastructures and at the same time equip us to face the challenge in view of the global patenting regime.

Studies on the highly endangered species

Similarly, genetic studies on highly endangered species such as swamp deer; Indian rhino, elephant, lion and tiger, etc. are of immediate importance. Overall genetic diversities amongst these animals may be ensured only when we understand the extent of prevalent genetic heterogeneity. A careful and concerted approach without disturbing them or their ecological niche will go a long way to secure a safe place for these natural *beauties* that have been poached in the name of *beast*.

Gene and genome mapping

As mentioned earlier, gene and genome mapping efforts should be undertaken without any further delay. This is crucial not only for developing synteny but also for precise localization of the genes on the chromosome. For *in situ* hybridization, conventional approach using radioisotopes or fluorescence *in situ* hybridization (FISH) may be used. The FISH approach has many other advantages over conventional *in situ* hybridization techniques.

One of the most reliable approaches that have gained momentum for genome mapping is the hybridization of contig(s) with monochromosome hybrids. Use of monochromosome hybrid assigns unequivocally the correct chromosomal position to a gene though the exact location would still involve FISH approach. Thus, a research center should be equipped to deal with not only established conventional *in situ* approach but also FISH and monochromosome hybridization facilities. Monochromosome hybridization involves generation of heterokaryons and systematic propagation of retained chromosome(s) in the background of other much frequently 'knocked-out' genome in the cell culture. The technology is well established and is routinely used for human system. However, for animal species, this has remained relatively unexplored. Reports are available on the formation of human/mouse heterokaryons and selective elimination of human chromosome in the background of mouse but no effort has been made, for example, for the formation of cell hybrid of buffalo and mouse or cow and rat. It is likely that we learn the very mechanism of elimination of a selective genome in the process of generating heterokaryons using cell lines from the farm animals. Thus, animal cell fusion (along the line of human system) and generation of heterokaryons would be yet another important experimental approach not only toward gene mapping but also for understanding the molecular events leading to the selective elimination of one genome in the background of another.

Gene regulation and expression

All the genes that express do not always result in translation of the peptides. There are regulatory mRNA transcripts that are equally important for the maintenance and upkeep of cell systems. Expression studies are conducted at the total RNA level employing northern blot analysis, by RT-PCR approach, RNase protection assay or nuclear run-on assay. Expression studies during the course of development in a species both in somatic and germline cells reveal the differential expression of a gene. If the level of expression varies consistently, from the animals of one population to that of others, it would reflect difference in the population structure. Thus, the population structures may also be uncovered with respect to coding and non-coding sequences. Finally, expression studies may also help in identifying the loci linked with genetic superiority of the animals such as resistance to diseases, tolerance to drought condition or less fodder requirement, etc.

Development of DNA-based genetic markers for species identification and conservation biology: The MASA strategy

An offshoot of the above studies would be the development of DNA-based genetic markers useful for addressing issues related to diversified aspects of animal biotechnology. Screening of the libraries would generate numerous clones representing expressed sequence (from cDNA library) or regulatory elements (from genomic library). Identification and characterization of such clones and their heterologous use will provide wealth of information in the context of comparative genome analysis. Similarly, identification and characterization of genes and their mutant alleles linked with satellite sequences, by minisatellite associated sequence amplification (MASA) would be equally rewarding. The animal genome like any other higher eukaryotes, is also endowed with different kinds of short tandem repeat motifs (STRs) in the range of 2–15 base pairs. Such STRs are represented several hundred to thousand times depending upon the overall rate of mutation[8,9] and the natural selection in favour of these satellites[35,36]. Minisatellite associated sequence amplification (MASA) studies using three STRs representing 16 base long consensus sequence of locus 33.15 and 15 and 18 base long 5′TGTC3′ repeat motifs revealed clad specific amplicons. MASA works on the basic principle of restriction fragment length polymorphism (RFLP) but the use of PCR, unlike RFLP, facilitates the quick detection of the signals obliterating the requirement of a large quantity of DNA and several days of auto-radiography. In case of actual experimental condition, at the outset, about 100 STRs from different satellite fractions in the range of 2–18 base residues can be used for conducting MASA with about 200–400 animal DNA samples representing well-identified populations. The MASA amplicons represented a number of functional genes associated with STR motifs. With a single primer, this approach does not amplify all the functional genes. However, by using different STRs derived from various satellite regions encompassing larger parts of the genome, a large number of functional genes may be amplified. In the event of alteration of a gene leading to its mutant allele, this can be identified by hybridizing MASA amplicons and assessing its allele length variation. It is envisaged that normal gene will show no change in the position of the amplicon(s) whereas the mutant alleles will show allele length variation. With this approach, a number of STRs may be screened and comprehensive gene-grids representing satellite sequences on one axis and genes linked with them on the other may be generated. It may be noted that with a single copy gene, it takes 10–15 days for obtaining hybridization signal. However, owing to a billion-fold amplification of the target substrate, the amplicons generated by MASA will produce signals in less than one hour. This approach will demonstrate most, if not all, of the coding sequences (functional genes) and their possible association with the satellite sequences. Thus, MASA approach provides ample working flexibility with respect to: (i) identification of a satellite-linked functional gene and its mutant alleles enabling us to uncover the population structure, and (ii) ascertaining the diseased loci to gain information about the susceptibility status of a population (population at risk). Subsequently, more refined gene grid, as mentioned earlier may be developed. A comparison of the data on allele length variation obtained from different populations would provide wealth of information on the population structure.

Man-power development

Simultaneously, serious attempts should also be made to train manpower (with infectious scientific enthusiasm) not only for delivering ‘today’ but also during the fast approaching next millennium. Despite many progresses made in the area of Science & Technology and availability of highly trained manpower, the potentials of our research scientists have remained underutilized. Thus, it is not only important to build *state of art* infrastructures that would be needed for conducting diversified experiments but also to expose researchers, particularly the new crop, to a global science culture. Similarly, strategies need to be evolved to initiate the process of developing an original thinker in a scientist than a mere research technician. In this context, an early identification of talent would prove to be as critical for success of science in the long run as asking a right kind of scientific question itself.

1. Ali, S., Verma, G. and Bala, S., *Anim. Genet.*, 1993, **24**, 199–202.
2. Ali, S., Ansari, S., Ehtesham, N. Z., Azfer, M. A., Homkar, U., Gopal, R. and Hasnain, S. E., *Gene*, 1998, **223**, 361–367.
3. Sulaiman, I. M. and Hasnain, S. E., *Theor. Appl. Genet.*, 1996, **93**, 91.
4. Sulaiman, I. M. and Hasnain, S. E., *Electrophoresis*, 1995, **16**, 1746.
5. Deka, R., Shriver, M. D., Yu, L. M., Ferrell, R. E. and Chakraborty, R., *Electrophoresis*, 1995, **16**, 1659–1664.
6. Mattapallil, M. J. and Ali, S., *DNA Cell Biol.*, (in press).
7. John, M. V. and Ali, S., *DNA Cell Biol.*, 1997, **16**, 369–378.

8. Weber, J. L. and Wong, C., *Hum. Mol. Genet.*, 1993, **2**, 1123–1128.
9. Shriver, M. D., Jin, L., Chakraborty, R. and Boerwinkle, E., *Genetics*, 1993, **134**, 983–993.
10. Di rienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M. and Freimer, N. B., *Proc. Natl. Acad. Sci. USA*, 1994, **91**, 3166–3170.
11. Ali, S. and Wallace, R. B., *Nucleic Acids Res.*, l988, **16**, 8487–8496.
12. Raina, A., Sulaiman, I. M., Das, P., Ehtesham, N. Z., Ali, S., Dogra, T. D. and Hasnain, S. E., *Gene*, 1996, **173**, 247–250.
13. Azfer, M. A., Bashamboo, A., Ahmed, N. and Ali, S., *J. Biosci.*, 1999, **24**, 101–107.
14. Rao, B. K., Sil, S. B. and Majumder, P. P., *J. Genet.*, 1997, **76**, 181–188.
15. Shriver, M. D., Jin, L., Ferrell, R. E. and Deka, R., *Genome Res.*, 1997, **7**, 586–591.
16. Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R. and Cavalli-Sforza, L. L., *Nature*, 1994, **368**, 455–457.
17. Ali, S. and Epplen, J. T., *Indian J. Biochem. Biophys.*, l991, **28**, 1–9.
18. Ehtesham, N. S., Das, A. K. and Hasnain, S. E., *Gene*, 1992, **111**, 261.
19. Ehtesham, N. Z. and Hasnain, S. E., *Adv. Forensic Haemogenet.*, 1992, **4**, 137.
20. Raina, A., Sulaiman, I. M., Ehtesham, N. Z., Das, P., Ali, S., Dogra, T. D. and Hasnain, S. E., *Gene*, 1996, **173**, 247.
21. Ehtesham, N. Z., Talwar, G. P., Ali, A. and Hasnain, S. E., *Indian J. Biochem. Biophys.*, 1990, **27**, 275.
22. Ehtesham, N. Z., Ma, D. P. and Hasnain, S. E., *Gene*, 1991, **98**, 301.
23. Ali, S., Azfer, M. A., Bashamboo, A., Mathur, P. K., Malik, P. K., Mathur, V. B., Raha, A. K. and Ansari, S., *Gene*, 1999, **228**, 33–42.
24. Dil-Afroze, Misra, A., Sulaiman, I. M., Sinha, S., Sarkar, C., Mahapatra, A. K. and Hasnain, S. E., *Gene*, 1998, **206**, 45.
25. Sinha, S., Dil-Afroze, Misra, A., Sulaiman, I. M., Sarkar, C., Mahapatra, A. K. and Hasnain, S. E., *FASEB J.*, 1997, **11**, 2313.
26. Mattapallil, M. J. and Ali, S., *Gene*, 1997, **206**, 209–214.
27. Takezaki, N. and Nei, M., *Genetics*, 1996, **144**, 389–399.
28. Chakraborty, R., Kimmel, M., Stivers, D. N., Davison, L. J. and Deka, R., *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 1041–1046.
29. Sulaiman, I. M., Ehtesham, N. Z. and Hasnain, S. E., *Gene*, 1995, **156**, 223.
30. John, M. V., Parwez, I., Sivaram, M. V. S., Mehta, S., Marwah, N. and Ali, S., *Gene*, 1996, **172**, 191–197.
31. Kimmel, M., Chakraborty, R., King, J. P., Bamshad, M., Watkins, W. S. and Jorde, L. B., *Genetics*, 1998, **148**, 1921–1930.
32. Ali, S., Verma, G. and Bala, S., *Mol. Cell. Probes*, 1992, **6**, 521–526.
33. Ali, S., Appa Rao, K. B. C. and Bala, S., *Assisted Reprod. Rev.*, 1993, **3**, 37–43.
34. Ali, S., Ravindranath, N., Jia, M. C., Musto, N. A., Tsujimura, T., Kitamura, Y. and Dym, M., *Biochem. Biophys. Res. Commun.*, 1996, **218**, 104–112.
35. Deka, R., Shriver, M. D., Yu, L. M., Jin, L., Aston, C. E., Chakraborty, R. and Ferrell, R. E., *Genomics*, 1994, **22**, 226–230.
36. Ali, S., Müller, C. R. and Epplen, J. T., *Hum. Genet.*, l986, **74**, 239–243.

Received 4 June 1999; accepted 21 July 1999

Optimality principles in evolutionary genetics

Narayan Behera

A population evolves due to changes in its gene frequencies arising due to mutations, natural selection, random genetic drifts and migrations. Svirezhev introduced an integral variational principle, in analogy with the least-action principle of classical mechanics, by defining a Lagrangian which remained stationary on the trajectory followed by the population undergoing selection. This principle can also be extended to multiple loci in some simple cases. However, in a two-locus model or more general models, there is no straightforward extension of this principle if linkage and epistasis are present. The local optimality principle can be geometrically formulated in a Riemannian metric space of gene frequencies so that, under evolutionary pressures, the population trajectory moves in that direction along which the increment of the mean fitness is maximum.

The causes of evolution

Gene is a physical entity, transmitted from parent to offspring, that influences a hereditary trait. From a biochemical point of view, genes are fundamental units of genetic information that correspond to the sequence of nucleotides in a segment of DNA. The position of a particular gene in a chromosome is called a locus. Genes can exist in different forms or states. For example, a gene for hemoglobin may exist in a normal form or in any one of a number of alternative forms that result in hemoglobin molecules that are more or less abnormal. These alternative forms of a gene are called alleles. The **genepool** of a population is defined as the collection of genes belonging to all the members of a population taken together. **Genotype** is the hereditary or genetic constitution of an individual. Genotype thus refers to the particular alleles present in an organism at all loci that affect the trait in question. The sum total of observable structural and functional properties of an organism, which is the product of the interaction between the genotype and the environment, is the **phenotype** of the individual. **Epistasis** is the mode of interaction among nonallelic genes which results in the phenotype of an organism (of a given genotype) that is not mere 'sum' of the phenotypes caused by genotypes at each locus taken separately. **Linkage** is the association in the inheritance of the genes that are located in a single chromosome (i.e. linkage group), and are termed the linked ones. Breakage of the joint transmission of the linked gene occurs as a result of recombination. The formulation of a problem in population genetics involves first assuring that some trait is the expression of a specific genetic mechanism. The net numbers of surviving progenies that an organism (genotype) leaves behind, to the next generation (relative to that of other genotypes), is defined as its **fitness**. Fitness is usually measured in terms of malthusian parameter, which is the difference between the birth rate and death rate.

In a sexually reproducing population, the genes get scrambled in each generation owing to segregation and recombination, the genotypes constantly changing as individuals die and new ones are born. The unique expression of a particular trait of a genotype depends on the manner in which the alleles of a gene interact during the development of an organism. For the expression of any of the alleles of one gene, dominance refers to the concealment of the presence of one allele by the strong phenotypic effect of another. The basic mechanisms by which genes are passed on from one generation to the next were discovered by Mendel in 1865 (and were rediscovered by genetics experiments in the beginning of this century). His law of segregation states that every somatic cell of an individual carries a pair of genes for each character; the members of each pair separate during meiosis so that each gamete carries

Narayan Behera is in Evolutionary Population Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands; and also at Theoretical Sciences Unit, Jawaharlal Nehru Center for Advanced Scientific Research, Jakkur, Bangalore 560 012, India.

one gene of each pair, and gametes pair at random to form the genotype of the offspring. The law of independent assortment says that each member of any pair of alleles is equally likely to be combined with either member of another pair of alleles, since they associate independently.

Genotypes undergo shuffling during gamete formation owing to segregation and recombination, and are re-assembled in each generation by the process of fertilization. Due to random mating, the probability of choosing a particular genotype for a mate is equal to the relative frequency of that genotype in the population. In the absence of factors that change the gene frequencies (described below) coupled with random mating, the population will have constant gene and genotype frequencies from generation to generation. This is called the Hardy–Weinberg principle, which is true only for an infinitely large population. Many populations in nature have births and deaths occurring more or less continuously, with both reproduction and mortality at various ages.

Thus, the changes in the gene frequencies in a population is the process of **evolution**. Evolution is caused by the following four basic forces acting on the population. (i) *Mutation*: Mutation occurs by ‘chance’ due to error in copying of chromosomes. It is a statistical fluctuation undergoing in nature. One allelic form changes into another under mutation. Mutation is the ultimate source of new or novel genes and it also prevents old alleles from ever being entirely eliminated. Recurrent mutation is fundamental to the process of evolution because it helps to maintain a supply of genetic variation for selection to act on. (ii) *Natural selection*: The fitter genotypes survive to produce more progeny for the next generation compared to the less fit genotypes which cannot leave their genes in the genepool. As long as the alleles which are present in the population do not all have the same fitness values, natural selection will change the state of the genepool. Selection occurs when one genotype leaves a different number of progeny than another. This may happen because of differences in survival, in mating or in fertility. This is the **Darwinian paradigm**. The idea of natural selection is an extension of the idea of artificial selection (as practiced by animal and plant breeders) except for the fact that in the former no conscious agent is involved to bring about selection. The spirit of the Darwinian theory of natural selection is that natural selection uses the genetic variation existing in a population to produce individuals that are better adapted to their environment. Through natural selection, alleles that enhance survival and reproduction increase gradually in frequency from generation to generation, and the population becomes progressively better able to survive and reproduce in the environment. A general result of natural selection is the progressive genetic improvement in the population which constitutes the process of evolutionary adaptation. (iii) *Genetic drift*: Genetic drift is the change in the general structure due to errors or ‘noise’ that creeps into the transmission of the genepool from generation to generation. Random sampling of alleles in a finite population results in chance changes in allele frequencies. The random sampling from the gamete pool means that some alleles may be over represented in current generation relative to their frequencies in previous generation and some alleles may be underrepresented. An allele can get fixed in the population after the passage of few generations by pure chance. (iv) *Migration*: Many species occupy a rather broad area and they may also be subdivided into several more or less separate populations. Most subpopulations are connected to one another through recurrent interchange of migrants. The population subdivision prevents random mating from occurring. However, the migration among the subpopulations can result in change in the gene frequencies.

Selection changes the variance in a given population. Whether the variance increases or decreases depends on the gene frequencies, dominance, epistasis, linkage and the mating system. Natural selection is based on total fitness, and the effect on any measurable trait depends on the relation between that trait and the expectation of surviving and reproducing. Natural selection, like classical mechanics, has both static and dynamic aspects. The former involves the relatively stable situation that results from the balance of various opposing forces – mutations, selection, migrations, and random fluctuations.

Fundamental theorem of natural selection

The variance of any trait in a population is in general determined partly by genetic factors and partly by environmental factors. We are primarily interested in that part of the variance which is determined solely by genetic differences among individuals in the population. Genic values are measures of the contribution of each gene, i.e. averaged over all genotypes into which this gene enters, and are measured as a deviation from the mean. Fisher enunciated his Fundamental Theorem of Natural Selection as: the rate of increase of mean fitness of a population at any time is equal to its genic variance in fitness at that time[1]. This holds true when the fitness depends on the genes at one locus only in a randomly mating system. The theorem accounts for the effect of gene frequency changes alone, isolated from the other things that are happening. **Fitness** measures the evolutionary improvement of a population brought about by the changes in the gene frequencies. Natural selection tends to preserve those genes which increase the fitness of their carriers. Intuition tells us that the rate at which selection changes the fitness will be related to the variability of the population. Under natural selection, the population follows a trajectory in the space of gene frequency such that the fitness of the population increases at a rate equal to the genic variance. This is the first instance of an optimality principle in population genetics. The Fundamental Theorem is compared with the second law of thermodynamics where entropy always increases in a physical process. The status of Fisher's theorem has long been controversial in evolu-tionary theory. It is well known that it is not exactly true in any but the simplest models of population genetics. When fitness depends on the presence of genes at more than one locus, the total change in mean fitness can be negative, even under random mating system. This essentially happens due to recombination and nonadditive gene effects. If it is reinterpreted as an approximate statement, i.e. if it is taken to state that the rate of change of fitness is approximately equal to the genic variance, then its range of applicability is significantly enhanced. Thus, the new interpretation of Fisher's theorem, put forward by Ewens, focuses on the partial change in the mean fitness, rather than the total change in the fitness[2]. Fisher attributed this component of the change in mean fitness to the changes in gene frequency caused by natural selection, as opposed to the change due to the environment. This partial change is the genic variance which is irrespective of the mating system. Furthermore, this result holds for multiple loci with multiple alleles at each locus, and for arbitrary recombination structure between the various loci. Thus, the partial increase in mean fitness is always nonnegative, and is due to the average effects of the alleles which form the basic components in the new interpretation. Fisher's theorem has an interesting analog for the rate of change of a character, which is correlated to its fitness, wherein one replaces 'genic variance' by 'genic covariance' in the statement of Fisher's theorem to obtain an interesting true generalization[3,4]. Thus, using the approximate form of Fisher's Theorem, one can calculate or measure the genic variance and, subsequently, predict that the rate of change of a population's fitness will be approximately equal to it. Fisher dealt with a continuous population and was not clear whether the theorem could be applied to discrete generations as well. Kimura was the first to add explicit terms into the theorem to account for environmental changes (or other causes of changes in the genotypic fitnesses)[5]. His treatment has been widely regarded as clarifying and extending Fisher's Fundamental Theorem. While the Fundamental Theorem states that the average fitness of a population increases as long as the population is not in equilibrium, Kimura's maximum principle states that the change in gene frequencies occurs in such a way that the increase in average fitness is maximal.

Price's analysis of Fundamental Theorem recognizes the fact that the theorem refers to the partial change[6]. Finally, Ewens' clearer formulation of the theorem led to its wider acceptance[2]. Ewens believed that Fisher was interested in fitness changes only through changes in the gene frequency and not through changes in the expected value of the fitness. The theorem, although exact, measures only one component of the fitness change. So the modern interpretation of the Fundamental Theorem (following Ewens and Price) is: 'the rate of increase in the mean fitness of any population at any time ascribable to natural selection acting through changes in gene frequencies is exactly equal to genic variance in the fitness at that time'. However, while Ewens' new formulation does not require any specific mating systems, and holds good for both discrete and continuous time, Fisher's result is true for random mating only. Thus, it is clear that the calculation of the partial change is exact and involves no approximations. This supports the argument that Fisher himself viewed the theorem as exact (for some background discussion, see ref. 7).

Price was perhaps correct in saying that Fisher viewed the partial changes in mean fitness as those due to single-locus gene-frequency changes. These changes essentially form the basis of evolution according to Fisher (in contrast, for example, to the evolution of co-adapted gene complexes favoured by Wright[8]). This may be the central tenet of the Fundamental Theorem. Nagylaki has shown that, under most circumstances, the terms additional to the genic variance are of a smaller order than the genic variance; the major exception occurs when the population is near an equilibrium[9]. Nagylaki believed that the biological significance of the partial change in the mean fitness remains to be demonstrated[10]. Ewens also recorded a negative assessment of the theorem as a biological statement [2]. Even if the Fundamental Theorem is inexact and incomplete, it captures the essence of the way selection works, and encapsulates a great deal of evolutionary insight into a simple expression[11]. The presence or absence of dominance, in the effect of genes, can play an important role for the optimality principle to hold true[12].

Global optimization

Hamilton's principle in classical mechanics can be stated as follows: Of all the possible paths along which a dynamical system may move from one point to another within a specified time interval (consistent with any constraints), the actual path followed is that which minimizes the time integral of the difference between the kinetic and potential energies. The variational statement of the principle states that the integral of the difference of kinetic and potential energies is an extremum, and not necessarily a minimum. But in almost all important applications in dynamics, the minimum condition occurs.

The intuition that leads to the expectation that the optimality principles must exist is, that if selection acts to increase a well-defined quantity (called 'fitness') over a period of time during which selection acts, then some quantity correlated with fitness ought to be maximized. In a standard, continuous time, one-locus selection model with random mating and constant (frequency- and time-independent) fitness, Svirezhev showed that for the trajectory followed by the population in its configuration in space, the integral of the 'Lagrangian' is:

$$\frac{1}{8}\left(\sum_i (\dot{p}_i^2/p_i) + \sum_i p_i(m_i - \overline{m})^2\right),$$

(where p_i is the frequency of the ith allele, m_i its (marginal) fitness, m is the mean fitness of the population and $\dot{p}_i = \mathrm{d}p_i/\mathrm{d}t$) is stationary[13]. This principle is a very close analog to Hamilton's principle in classical mechanics where the Langrangian is equal to T (the kinetic energy) – V (the potential energy) of a system: this integral is known as the action. This is reminiscent of the principle of least action and Hamilton's principle in classical mechanics whereby a minimization process leads to Newton's laws of motion (see Lanczos[14] for a detailed discussion of the history and the limitation of this principle).

The range of validity of Svirezhev's principle has recently been explored for multilocus genetic system in a systematic way[15,16]. To be precise, that principle will be taken to state that a population undergoing selection follows a trajectory that keeps the integral of a Lagrangian function stationary, where the Lagrangian has the form indicated above. This particular form of the Lagrangian, that Svirezhev had introduced, is interesting because, leaving out the factor of 1/8, it is exactly equal to the variance. The second term is obviously half the variance (half because of diploidy) and the first term is also half the variance because,

$$\dot{p}_i = p_i(m_i - \overline{m}).$$

However, these statements are true only on the natural-selection trajectory. The potential utility of this principle is clearly the same as Fisher's: one writes down the formula for the genic variance, and then expects the population to follow the trajectory where (to some approximation) the integral of one-eighth of the variance remains stationary. The general strategy followed throughout this paper is therefore: (i) the equations governing the change of allelic or gametic frequencies over time are written down (which are derived using the standard techniques of population genetics), (ii) using Svirezhev's principle, a Lagrangian function is written down, (iii) the Euler–Lagrange equations corresponding to the Lagrangian are derived, using the method of Lagrange multipliers to incorporate the known constraints. This is the stage at which the calculus of variations plays a role. That the integral of the Lagrangian is stationary (over a trajectory) is equivalent to the condition that the Euler– Lagrange equations hold for the motion, (iv) these Euler–Lagrange equations are solved to obtain the equations of motion for the system, and (v) this solution is compared to the equations that were obtained using the standard techniques. When linkage disequilibrium is zero, the rate of change of allelic frequency can be correlated with Malthusian parameters (hence with experimentally observable quantities).

Genetic flexibility, for adaptation to a fluctuating environment, demands recombination to generate adaptive combinations different from those previously selected. The roles of linkage and epistasis are very important in the context of genomic evolution. In the above analysis, the rate of change of allelic frequency is used, for mathematical convenience, to prove the variational principle. With linkage and epistasis, the rates of change of allele frequencies no longer suffice to determine the genotypic dynamics of the population, instead gametic (or haplotype) frequences must be used. Svirezhev's principle does not hold good even in simple models when linkage and epistasis are incorporated[16].

Population trajectories in a metric space

All trajectories of the population equations have been examined only in the Euclidean space. In the Euclidean space, the trajectories of the population equations do not provide the steepest ascent for the mean fitness. Hence, nothing can be said about the dynamics of the movement of the population towards the stable equilibrium condition. However, in a suitable Riemannian metric space, the trajectories can be constructed as a gradient of mean fitness, so that, under evolutionary pressure, the population selects that direction out of all the possible directions along which the increment of the mean fitness is maximum. It would be better if these trajectories possessed some external features, for instance if they were trajectories of a vector field of a gradient of mean fitness. The local principle of optimality can be formulated as follows. If we examine the motion of a population in a certain Riemannian metric space of gene frequencies with the basic form of the matrix tensor g_{ij} defined as:

$$g_{ij} = \tfrac{1}{2}\frac{p_i\delta_{ij} + p_j\delta_{ji}}{p_i p_j},$$

where δ_{ij} is the delta of Kroneker, then all trajectories of its motion in this space lie on the unit sphere. At any moment in time, from all possible directions on this sphere of the population, such a path is chosen that the increase of its average fitness is a maximum, and the population moves at a rate proportional to a certain degree of the genetic variation in the population. This rate V equals[13]

$$|V| = \sqrt{\Sigma\, p_i\,(m_i - \overline{m})^2}\,.$$

In the n-locus system (when all possible linkage disequilibrium are zero), the corresponding surface on which the phase point moves is the product space of n higher-dimensional unit spheres: When linkage and epistasis are present, the metric tensor will be dependent on fitness parameters in addition to gametic frequencies; and in a two-locus two-allele genetic system, the population trajectories will lie on a three-dimensional surface with a complicated geometry[16].

The particular form of metric tensor, used by Svirezhev for mathematical convenience, was given suitable biological justification[17]. Thus, overall action of natural selection minimizes the genetic distance between parental and daughter gene frequency values such that the partial increase in mean fitness is the natural selection value. This genetic distance has the same quadratic functional form as the metric. This new interpretation is based on the concept of the average effect of an allele. The gradient pattern of genetic process of variations in allele frequencies under selection pressure as a motion in a Riemannian space as well as of the properties of this motion in the case of several loci, was considered[18]. Therefore while examining the trajectory of a population as the trajectory of a kind of mechanical motion in a special chosen space, we cannot link its geometry to the characteristics of selection. Instead, this geometry is determined by a system of interbreeding and by-laws of heredity. It is interesting therefore to look at the trajectory of a population in such a space, the geometry of which would immediately be linked to characteristics of selection, for example in the metric space with the basic form:

$$ds^2 = a_{ij}\,dx^i dx^j,$$

$i, j = 1, 2, \ldots n$, where a_{ij} are the coefficients of the relative viability and dx^i is the infinitesimal change in the contravariant vector x^i, the components of which are equal to the allelic frequencies p_i. It is assumed that det // a_{ij} // $\neq 0$. Thus, in such a system, polymorphism is present, i.e. in the population 'the persistence of two or more alleles at the fixed locus in the population'. However, this method appears fruitful only for the analysis of the stationary condition existing in a polymorphic system[13].

Conclusion

Minimal principles in physics have a long history. The search for such principles is primarily motivated on the ground that nature always minimizes certain important quantities when a physical process takes place. In 1657, Fermat formulated such a principle by postulating that a light ray always travels from one point to another in a medium by a path that requires the least time. The principle of least action was mathematically formulated by Lagrange (1760) who asserted that dynamical motion takes place with minimum action. Hamilton (1835) defined the dynamical principle on which it is possible to base all of mechanics. Hamilton's principle allows us to calculate the equations of motion of a body completely without recourse to Newtonian theory. The equations of motion are obtained without the necessity of explicitly taking into account the external forces acting on the body. According to Hamilton's principle, in Nature, the motion of a body occurs to minimize the time integral of the difference between the kinetic and potential energies. Furthermore, using this approach, the connection between symmetry properties and the invariance of physical quantities can be suitably formulated. Hamilton's method is in essence a *posteriori*, because we know beforehand that a result equivalent to Newtonian equations must

be obtained which are correlated with experimental facts. Hamilton's principle, based on a single basic postulate, gives satisfying unification of many individual theories.
It is a goal of physical theories to describe observed phenomena mathematically with an economy of fundamental postulates.

The failure to extend Svirezhev's principle to models with linkage does not mean that no function exists such that the Euler–Lagrange Equations obtained using it (as the Lagrangian) would give rise to the appropriate dynamics. However, unless there exists some method by which the Lagrangian function can be obtained *without already knowing the dynamics*, the Lagrangian reformulation of these models is of little value. In physics, this is straightforward because the Lagrangian is equal to $T - V$, where T is the kinetic energy and V is the potential energy of the system. Similarly, in these population genetics models, without linkage or epistasis, the Lagrangian is simply one-eighth of the variance written in Svirezhev's form. The failure of the Svirezhev principle however means that this recipe for finding the Lagrangian no longer works. Moreover, there is no straightforward way to change the recipe, for instance, by the introduction of an additional additive term in the Lagrangian[16]. In the case of Fisher's Fundamental Theorem, the presence of linkage and epistasis can be taken into account (in two-locus continuous time models) by adding terms to the genic variance (see Crow and Kimura[19]). However, the results reported in ref. 16 show that such a strategy is not available for the Lagrangian formulation. This formulation cannot play a particularly useful heuristic role in any but the simplest situations where linkage disequilibrium is zero. The success of Svirezhev's principle in these cases suggests the following conjecture: this principle is applicable in exactly those circumstances where Fisher's Fundamental Theorem, in its exact form, i.e. the rate of change of the mean fitness is exactly equal to the genic variance, is true.

The formulation of the population trajectory as a trajectory of steepest ascent in a Riemannian metric space provides an insightful way of looking at evolution. The analysis of Svirezhev's variational principle in the population genetics models involving: (i) non-random mating, and (ii) multiple loci with non-zero linkage disequilibria are mathematically too cumbersome. In this work some analogies have been formulated between classical mechanics and certain problems in mathematical theory of population genetics.

1. Fisher, R. A., *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, 1930.
2. Ewens, W. J., *Theor. Pop. Biol.*, 1989, **36**, 167–180.
3. Crow, J. F. and Nagylaki, T., *Am. Nat.*, 1976, **110**, 207–213, 400.
4. Nagylaki, T., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 1910–1913.
5. Kimura, M., *Heredity*, 1958, **12**, 145–167.
6. Price, G. R., *Annu. Hum. Genet.*, 1972, **36**, 129–140.
7. Edwards, A. W. F., *Biol. Rev. Cambridge Philos. Soc.*, 1994, **69**, 443–474.
8. Wright, S., *Am. Nat.*, 1988, **131**, 115–123.
9. Nagylaki, T., *Genetics*, 1976, **83**, 583–600.
10. Nagylaki, T., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 2402–2406.
11. Crow, J. F., *Theor. Pop. Biol.*, 1990, **38**, 263–275.
12. Narain, P., *J. Genet.*, 1993, **72**, 59–71.
13. Svirezhev, Y. M., *Studies on Theoretical Genetics* (ed. Ratner, V. A.),, USSR Academy of Sciences, Novosibrisk, 1972, pp. 86–102.
14. Lanczos, C., *The Variational Principles of Mechanics*, University of Toronto Press, Toronto, 1959.
15. Behera, N., *J. Genet.*, 1995, **74**, 19–24.
16. Behera, N., *Bull. Math. Biol.*, 1996, **58**, 175–202.
17. Ewens, W. J., *Theor. Pop. Biol.*, 1992, **42**, 333–346.
18. Shahshahani, S., *Mem. Am. Math. Soc.*, 1979, **17**.
19. Crow, J. F. and Kimura, M., *An introduction to Population Genetics Theory*, Harper and Row, New York, 1970.

ACKNOWLEDGEMENT. I thank Sahotra Sarkar for helpful discussions.

Received 18 March 1998; revised accepted 16 September 1999