ABSTRACT

Human reproductive ecology is a relatively new subfield of human evolutionary biology focusing on the responsiveness of the human reproductive system to ecological variables. Many of the advances in human, and more recently primate, reproductive ecology concern the influence of energetics on the allocation of reproductive effort. This paper reviews eleven working hypotheses that have emerged from recent work in reproductive ecology that have potential bearing on the role of energetics in human evolution. Suggestions are made about the inferences that may connect this body of work to our efforts to reconstruct the forces that have shaped human biology over the course of our evolutionary history.


There is increasing interest in the role that energetics may have played in shaping important aspects of human evolution (Aiello and Key 2002; Leonard and Ulijaszek 2002). By energetics I mean those processes by which energy is captured from the environment for metabolic purposes and allocated to those purposes. Such a broad definition captures a great deal that has traditionally been a part of theorizing about human evolution, including diet and foraging behavior (Sorensen and Leonard 2001), digestive physiology (Milton 1987), and the energetic efficiency of locomotion (Hunt 1994). But it also includes the consequences of necessary trade-offs in energy allocation among competing physiological domains, such as physical growth, immune function, thermoregulation, subsistence work, and reproduction. From the perspective of life history theory, trade-offs between reproduction and survival (under which many of the other categories of metabolic allocation can be placed) are central to understanding evolutionary dynamics (Stearns 1992). Differential allocation to reproduction is known in this branch of evolutionary ecology as “reproductive effort.” The central premise of life history theory is that natural selection has acted on organisms to optimize the allocation of reproductive effort over the organism’s life span. This perspective has been remarkably successful in explaining the diversity of life histories observed in nature.

Of late there has been particular attention paid to the energetic aspects of human brain evolution (Aiello and Wheeler 1995; Leonard et al. 2003; 2007). The human brain is energetically very expensive, both to grow and to maintain. Meeting the energetic cost of such a large brain is a challenge that has both phylogenetic and ontogenetic implications. Phylogenetically, the question might be posed as, “How was the energy for such a significant new allocation made available in the budget of an organism that must be assumed to have already been allocating energy optimally?” Ontogenetically, the question may be posed as, “How is a human mother able to meet the allocation demand of her fetus and infant’s growing brain?”

Several hypotheses have been advanced in recent years to address the first question. Most of the hypotheses are based on an explicit awareness of the importance of energetic trade-offs. Thus, Aiello and Wheeler have proposed that reduction in gut size due to dietary changes in the human lineage may have lowered the maintenance costs of that tissue, freeing up metabolic energy for allocation to brain tissue (Aiello and Wheeler 1995). Leonard and Robinson make a similar argument, but point to a reduction in skeletal muscle mass rather than a reduction in gut tissue for the energy savings (Leonard et al. 2003). Wrangham et al. (1999) have suggested that the adoption of cooking may have lowered the metabolic costs associated with feeding and digestion, again freeing up energy for the brain. None of these hypotheses, however, speak explicitly to the ontogenetic question; instead, they focus on energetic allocation in adults.

Human reproductive ecology is a relatively new and growing field within human biology (Ellison 1994). Its focus is on the way in which evolution has shaped human reproductive physiology. Life history theory has been an important theoretical orientation for much of the work in this field as well. Energetic trade-offs between reproductive effort and other domains of physiology are a particular focus of investigation and theorizing (Ellison 2003). Recently, many of the paradigms that have been developed in the context of human reproductive ecology have been extended to non-human primates, and particularly to the reproductive ecology of wild apes (Emery Thompson 2005; Knott 2001).

There is a great deal in common between the perspective of human reproductive ecology and the concern for energetics among human paleontologists. But as yet the two areas of research have not intersected a great deal. The purpose of this paper is to help foster that intersection.
To do so, I will review the major working hypotheses that have arisen from contemporary work in human and primate reproductive ecology relating to the role of energetics in shaping reproduction. I will try to point out areas of continuing disagreement and controversy, as well as areas of prevailing consensus. This review will not be comprehensive, but will rather be goal directed, aimed at drawing together empirical and theoretical elements that can contribute to thinking about the role of energetics in human evolution. I will then offer some more speculative thoughts about the implications of integrating this information into schema of our evolutionary history.

**ENERGETICS AND THE ECOLOGY OF HUMAN REPRODUCTIVE EFFORT**

In presenting a summary of working hypotheses from human reproductive ecology I will use the term “energy availability” to refer to the availability of metabolic energy (as opposed, for example, to energy available in the environment), and in particular that fraction that might be allocated to reproduction. It is worth noting at the outset that energy availability can be limited either by constraints on energy input (including the mobilization of stored energy), or by constraints on energy expenditure. Input constraints are sometimes referred to as constraints on oxidizable substrates, the substrates from which biochemical energy in the form of high-energy phosphate bonds is derived (Wade and Schneider 1992). These substrates ultimately come from ingested food, but may be temporarily stored in complex carbohydrates (such as glycogen or starch), fatty acids, or structural proteins. Periods of low food intake can restrict energy availability, although the mobilization of stored energy can buffer this effect to some extent. The digestion of food and the mobilization of stored energy involve processes that are collectively known as catabolic, because they involve the breakdown of larger storage molecules to release oxidizable substrates.

Energy expenditure refers to the use of biochemical energy to do metabolic work (Ulitzske 1995). That expenditure can involve physical work in the form of locomotion and other physical activity based on muscle contraction, or it can involve anabolic or maintenance work within the organism. Anabolic processes are those that involve the production of larger molecules from smaller substrates, including the production of proteins, nucleic acids, fatty acids, complex carbohydrates, and more complex molecules constructed from these and other elements. Anabolic processes underlie physical growth, but they also are necessary for tissue maintenance and to balance the turnover of biologically active molecules. Secretory tissues are sites of high anabolic activity. The immune system involves dispersed populations of cells many of which undergo clonal proliferation and/or become secretory during periods of immune activation, both of these states involving anabolic processes. In addition to physical and anabolic work, a third important category of metabolic energy expenditure is in maintaining ionic gradients across cell membranes, particularly of nervous and muscle tissue, and other forms of active transport. Finally, heat production itself can constitute a category of metabolic energy expenditure as substrate molecules are oxidized in ways that are decoupled from energy capture in phosphate bonds, allowing dissipation of energy as heat. This can be an important category of energy expenditure under certain circumstances, subserved particularly by brown adipose tissue in infants and by the up-regulation of thyroid hormone production and subsequent partial decoupling of ATP production from oxidative metabolism as an adaptive response to cold stress in some populations (Leonard et al. 1999; Leonard et al. 2002; Snodgrass et al. 2005).

The following are current working hypotheses from the field of human reproductive ecology with implications for reconstructing human evolution. In some cases, particularly early on, necessary accounts of associated physiology will be included.

**HYPOTHESIS 1: THE FECUNDITY OF NON-PREGNANT, NON-LACTATING WOMEN IS SENSITIVE TO ENERGY AVAILABILITY**

Fecundity, in American usage, refers to the biological capacity to reproduce. One of the important contributions of reproductive ecology has been to identify ovarian function as an important component of female fecundity and to demonstrate the sensitivity of ovarian function to energy availability (Ellison 2001b). Ovarian function can be characterized in a number of ways, but most characterizations are reflections of the production of the major steroid hormones of the ovarian cycle: estradiol and progesterone (Figure 1). Estradiol is produced by developing follicles during the first half, or “follicular phase,” of the human menstrual cycle. It is also produced during the second half of the ovarian cycle, but its functional significance in that phase is unclear. Estradiol levels not only stimulate follicular development and oocyte maturation, they also stimulate proliferation of the endometrial lining of the uterus. If ovulation occurs, the ruptured follicle is normally transformed into a corpus luteum, the major steroid product of which is progesterone. Progesterone supports the maintenance of the uterine endometrium, which stops proliferating and becomes an active secretory tissue in preparation for the implantation of an embryo during what is known as the “luteal phase” of the ovarian cycle. If pregnancy does not occur, or if implantation is not successful, the corpus luteum regresses after about two weeks and the endometrial lining of the uterus is sloughed off in menstruation (Strauss and Williams 2004).

The frequency, timing, and duration of menstrual bleeding can provide one source of information about ovarian function. Infrequent menstruation almost always signifies infrequent ovulation, and the prolonged absence of menstruation indicates the absence of ovulation. Infrequent or absent ovulation necessarily reduces fecundity compared to regular ovulation. Highly variable menstrual patterns within individual women also are associated with low probability of conception compared to regular patterns in some studies (Kolstad et al. 1999). Irregular and infrequent menstruation is characteristic of the first years after men-
arche and the years preceding menopause in most women, and these are typically periods of relatively lower fecundity (van Zonneveld et al. 2003; Venturoli et al. 1989). But irregular and infrequent menstruation also is associated with restricted energy availability. High energy expenditure in the form of exercise has a well-documented association with infrequent and irregular menstruation (Broocks et al. 1990; De Souza 2003; Loucks et al. 1992), and menstrual disruption can be experimentally induced by exercise in randomized experimental designs (Bullen et al. 1985). Caloric intake restriction, weight loss, and emaciation, if severe enough, also are associated with menstrual disruption and even with complete cessation of menstrual cycling, whether experimentally induced, voluntarily pursued, or as the consequence of external constraint (Pirke et al. 1987; Pirke et al. 1985; Schneider 2004; Schweiger et al. 1987). The effects of restricted energy availability on menstrual function appear to be readily reversible when energetic constraints are relaxed or when weight is regained (Bullen et al. 1985; Prior et al. 1982). Thus rather than representing some lasting “damage” to a woman’s reproductive capacity, the correlation of menstrual pattern with energy availability appears to reflect a facultative ovarian response.

But menstrual patterns alone do not reveal the full range of variation in ovarian function. Regular ovarian cycles can vary in their levels and patterns of ovarian steroid production as well, both within and between women. This variation is much more difficult to study, however. Either frequent steroid measurements or some other quantitative index of steroid production is required. Regular (e.g., daily) venous blood sampling is generally too invasive and burdensome on subjects to be widely adopted as a research protocol. Döring used regular measurement of basal body temperature to indirectly reflect the timing of ovulation and duration of the luteal phase of the ovarian cycle in 3,264 person-months of observation obtained from 481 women (Döring 1969). His results documented the fact that a variable percentage of menstrual cycles that are outwardly normal in pattern are anovulatory, and another variable percentage is characterized by short luteal phases. Because short luteal phases often are associated with long preceding follicular phases, overall cycle length may appear normal. Yet there is good reason to believe that short luteal phases indicate cycles of diminished fecundity (Daly 1991; McNeely and Soules 1988; Soules et al. 1989). Certainly anovulatory cycles are infecund. Without an ovulated cycle, fertilization cannot occur.

The development of less invasive methods for monitoring ovarian steroid levels, in particular measurements made in samples of saliva or urine or in finger-prick blood

Figure 1. Average (pmol/L± standard error) daily concentrations for salivary estradiol and progesterone from 92 ovarian cycles collected by 24 Boston women (Lipson and Ellison 1996).
samples, have provided an important window into ovarian function that is not reflected in variation in menstrual pattern (Campbell 1994; Ellison 1988; Worthman and Stallings 1997). Studies based on these media can provide finer-grained information on ovarian function within individual subjects and allow for more extensive comparisons of variation within and between women. In addition, because protocols have been developed for the collection and preservation of samples under remote field conditions, these methods also have allowed for the study of ovarian function across a broad range of populations and environments. With these techniques it has become apparent that ovarian cycles can vary considerably in their steroid production even when menstrual patterns are constant.

A “continuum of ovarian function” has been suggested as a unifying framework for organizing the range of variability that has now been documented (Ellison 1990; Prior 1985) (Figure 2). This continuum extends from “textbook” ovarian cycles with high steroid profiles and regular menstrual patterns, though cycles with quantitatively lower steroid profiles and possibly altered ratios of follicular phase to luteal phase length, to anovulatory but still regular menstrual cycles, to cycles with variable menstrual frequency and duration, to the extended absence of menstruation. The appearance of menstrual pattern irregularity represents a normal “clinical horizon” for the recognition of variance in ovarian function. Variation in ovarian function below this horizon is subtler, requiring the monitoring of hormonal levels for its recognition.

The continuum of ovarian function is also thought to represent a continuum of fecundity. Variation above the clinical horizon—variation in menstrual pattern—is clearly linked to variation in fecundity, as previously noted. Variation below the clinical horizon—quantitative variation in steroid levels within menstrual cycles—also is associated with variation in fecundity by a number of different lines of evidence. One line of evidence is physiological, with many observations made in the course of developing and perfecting techniques of assisted reproduction. The level of follicular estradiol production is associated with the size of a follicle (Artini et al. 1994; van Dessel et al. 1996; Xia and Younglai 2000). Larger follicles produce oocytes that are more likely to be fertilized in vitro and that produce embryos of higher quality (which are more likely to implant successfully when reintroduced into the uterus) (Andersen 1993; Dubey et al. 1995; Miller et al. 1996). This effect of

---

**Figure 2.** The continuum of ovarian function. In this case the continuum has been presented with low fecundity at the top and high fecundity at the bottom of the figure to metaphorically correspond to conditions of amenorrhea and oligomenorrhea being “above” the clinical horizon. Detailed explanation provided in the text.
higher estradiol levels on the quality of individual oocytes is in part mediated by the effect of estradiol in promoting aspects of nuclear and cytoplasmic maturation in the oocyte (Driancourt and Thuel 1998). Follicular estradiol levels may also reflect follicular responsiveness to gonadotropin stimulation and thus may also be correlated with aspects of follicular quality that are independent of steroid action (Strauss and Williams 2004).

A second line of evidence is clinical, deriving from therapeutic effects of the manipulation of steroid levels. Luteal phase deficiency is a condition characterized by either low levels or short durations of progesterone secretion in the luteal phase of the cycle, or both (Lenton et al. 1984; McNeely and Soules 1988). It is associated with low fecundity in some women, and progesterone supplementation is an effective treatment in these cases (Daly 1991). Poor luteal function may in turn be a consequence of poor follicular development in some cases, so that the effects of low estradiol levels and low or short periods of progesterone elevation may be synergistic (DíZerega and Hodgjen 1981a, b).

A third line of evidence is epidemiological. Two studies have documented differences in steroid profiles in naturally occurring conception versus non-conception cycles in women attempting to conceive, one based on measurements of salivary steroid levels (Lipson and Ellison 1996), the other based on measurements of urinary steroid levels (Venners et al. 2006). In each case, the study design controlled for exposure to intercourse in the fertile period of the cycle. Both studies found that higher levels of follicular estradiol are significantly associated with higher probability of conception, both between and within women. Evidence for an effect of luteal progesterone levels on fecundity is more problematic to infer from comparisons of this nature, however, because the very presence of an embryo is known to affect luteal progesterone levels (Baird et al. 1997; Lu et al. 1999). Thus it is difficult to know what the progesterone level in a conception cycle would have been if the conception had not occurred, in order to compare it with that in a non-conception cycle.

Obesity is the one condition under which the direction of the relationship between energy availability and ovarian function can reverse. Rather than being associated with high levels of ovarian function, obesity can be associated with anovulation and menstrual irregularity (Diamanti-Kandarakis and Bergjke 2001; Franks et al. 1996; Norman and Clark 1998; Pasquali and Gambineri 2006; Rich-Edwards et al. 2002). There are a number of pathways that contribute to this association, including inappropriate negative feedback from excessive non-ovarian production of estrogens in fat tissue and high levels of androgen production. As a consequence, weight loss or exercise often can improve the ovarian function of obese women (Norman and Clark 1998; Vitzthum et al. 2004). While this is an important aspect of reproductive ecology for contemporary populations, I will assume, like most other anthropologists, that it has little significance for our understanding of human evolution, since I assume obesity has been a rare condition in our evolutionary past. However, as will become apparent, I do believe that our ability to store relatively great amounts of fat is evolutionarily significant.

Evidence from a number of different directions thus indicates that the continuum of ovarian function is also a continuum of fecundity. But two additional important points should be made here. First, the evidence for variation in fecundity associated with variation in ovarian function is restricted to comparisons within women and between women within populations. Vitzthum has stressed that variation in average ovarian steroid profiles between populations has not been reliably linked to between population variation in average fecundity (Vitzthum 2001; Vitzthum et al. 2004). Exploration of that issue will be difficult, since it will require careful control of exposure to intercourse across populations. In relation to this question, however, a second point must be made. The fact that an individual is characterized by low steroid profiles does not imply that the fecundity of that individual has dropped to zero, only that the probability of successful conception per cycle exposed to intercourse is lower than in a case with higher steroid profiles. If an individual has hormonal profiles that are chronically low, conception should still occur eventually. If we were to compare the average steroid profile in conception and non-conception cycles for this individual we would find they were the same. Variation in steroid profiles can only be associated with variation in conception success to the extent that it occurs.

This point is important because it is sometimes misleadingly stated that the high fertility of many traditional populations, despite evidence of low ovarian steroid profiles, is evidence against the idea that ovarian function affects female fecundity (Vitzthum et al. 2004). We must remember that, in the modern world, fertility (the actual production of offspring) and fecundity (the biological capacity to produce offspring) can be quite disjoint due to the availability of artificial contraception. Higher fertility in Eritrea than in Italy does not necessarily mean that Eritreans have, on average, higher fecundity than Italians. It may simply mean that Italians use contraception more often or more effectively than Eritreans, perhaps because of a greater desire for children on the part of the Eritreans. The working hypothesis from reproductive ecology is that female fecundity, not female fertility, scales with ovarian function. The only way to determine whether Eritrean or Italian women have higher fecundity would be to compare the waiting time to conception in women from both populations who are not using contraception and who have comparable exposure to intercourse.

By combining evidence based on changes in ovarian steroid levels with evidence based on changes in menstrual patterns, support for the hypothesis that female fecundity is sensitive to energy availability becomes even stronger. Even modest weight loss in normal weight women has been associated with reduction is both progesterone and estradiol levels in US and German women (Henley and Vaitukaitis 1985; Insler 1992; Lager and Ellison 1990; Lipson and Ellison 1996; Pirke et al. 1985; Schweiger et al. 1987). Higher levels of energy expenditure, whether from exercise or daily
work, have likewise been associated with reduced ovarian steroid profiles in women from a number of Western populations (Ellison and Lager 1985, 1986; Insler 1992; Jasienska and Ellison 1998; Jasienska et al. 2006c). Even when weight is not changing appreciably, restrained eating (eating below appetite) is associated with lower ovarian steroid levels as well as with elevated cortisol, an adrenocortical hormone that stimulates the catabolic release of stored energy (Barr et al. 1994; Berga et al. 2003; Warren et al. 1999). There is also evidence that low levels of stored fat and lipid profiles, indicative of reduced energy availability, are associated with reduced ovarian steroid profiles (Bruning et al. 1992; Furberg et al. 2005). Similar associations of negative energy balance (weight loss), elevated energy expenditure, and low energy status (fat reserves) with lower profiles of ovarian steroids among otherwise regularly cycling women have been observed in non-Western populations as well (Ellison et al. 1993; Panter-Brick et al. 1993; Vitzthum et al. 2002). It is notable that in many of these cases, increases in menstrual irregularity also occur, and among the women who continue to menstruate, a number have steroid profiles that suggest anovulatory cycles. Changes in menstrual regularity, ovulatory frequency, and ovarian steroid profiles usually go together. It is also worth noting that most of the studies of weight loss and energy expenditure have included within-woman comparisons, greatly reducing the impact of confounding factors such as age or ethnicity.

It is useful to think of a population of women as being distributed along the continuum of ovarian function at any point in time or under any specified set of conditions (Figure 3). Even under the most favorable conditions, some women may have low steroid profiles, some may have anovulatory cycles, and some may even have irregular cycles or be amenorrheic. Heterogeneity of this kind, independent of energy, is assumed. The weight of the evidence suggests that as energy availability becomes more limited, the distribution shifts upward on the continuum (as it is represented here). When the constraint on energy availability is modest, the greatest change in the frequency distribution appears to be in those states below the clinical horizon; that is, in states that are not apparent to a woman or her doctor on the basis of simple menstrual characteristics. As the constraints become more severe, changes in the frequency of states above the clinical horizon, including changes in menstrual periodicity, become more prevalent. Women also are expected to vary in the degree of their response to energy availability, and some women may remain unaffected under conditions that cause significant reductions in ovarian function in others. At the population level, however, the shift in distribution of individual women to lower levels of ovarian function will be associated with lower average fecundity.

The individual heterogeneity in ovarian function observable under any specific energetic conditions and the
heterogeneity in responsiveness to energy availability raise questions of functionality and adaptation (Ellison and Jasienska 2007; Jasienska et al. 2006a; Lipson 2001; Vitzthum 2001). Is it functional, in the sense of raising individual fitness, to shift to a lower level of fecundity under conditions of restricted energy availability? Or is it a manifestation of dysfunction? Are those individuals most fit whose ovarian function is unaffected by energetic conditions that cause reductions in ovarian function among others? If so, how do we account for the fact that ovarian function is not better buffered against variation in energy availability? If not, what accounts for the variation in individual sensitivity? Or does the heterogeneity in the distribution of ovarian function states and responses have other causes that, were they known, would resolve these issues? These are important questions that are difficult to answer empirically. My own view, and that of many others, is that ovarian responsiveness to energy availability is adaptive, helping to optimize the allocation of reproductive effort by adjusting the probability of conception to energetic conditions that are most favorable to reproductive success (Ellison 1990, 2001b, 2003). The reasons for that view will become more apparent as we consider additional hypotheses and the evidence that supports them. But it should also be noted that there are differences of opinion and interpretation even among those who adopt an adaptationist perspective to human reproductive ecology.

HYPOTHESIS 2: FEMALE FECUNDITY IN CHIMPANZEEs AND ORANGUTANs IS ALSO SENSITIVE TO ENERGY AVAILABILITY

Although the data for apes are more limited than for humans, there is accumulating evidence that ovarian function is responsive to energetics in both chimpanzees (Pan troglodytes) (Emery Thompson 2005) and orangutans (Pongo pygmaeus) (Knott 2001). Knott developed methods for collecting urine samples from wild orangutans in Borneo that have opened up the field of primate reproductive ecology to a wider range of research opportunities in much the same way that saliva and urine sampling earlier opened up the field of human reproductive ecology (Knott 1997, 2005). The samples sizes available from studies of wild ape populations are small, but significant results are still obtainable. The Bornean rainforest in which Knott conducted her studies is characterized by infrequent and irregular pulses of food availability for the orangutan population due to the masting habit of the dominant fruit trees—periods of high fruit availability in some seasons contrast with periods of dramatic fruit shortage at other times. The average caloric intake of female orangutans can vary 9-fold, from over 9,000 kcal/day to less than 1,000 kcal/day over the course of a few months as fruit availability goes from one extreme to the other (Knott 2001). During the periods of low caloric intake significant levels of ketones appear in the urine, indicating rapid mobilization of fat reserves (Knott 1998). Average estrogen levels in the urine also decline by more than 30% going from periods of high to low fruit availability (Knott 2001). Although orangutans have no readily visible signs of ovarian cyclicity such as sexual swellings or copious menstruation, mating activity is much higher during periods of high fruit availability than periods of low fruit availability (Knott 2001). Thus, the evidence so far indicates that ovarian function in orangutans is sensitive to environmental fluctuations in energy availability.

Similar evidence is now accumulating from wild chimpanzees. In a comparison of 11 conception cycles with 40 non-conception cycles from Gombe, Budongo, and Kibale, both urinary estrogen and urinary progestagen levels were significantly higher in the conception cycles (Emery Thompson 2005). For the estrogen levels this was true both in the swelling phase, before conception occurred, and in the post-swelling phase. In Budongo, urinary estrogen levels of individual females are observed to correlate positively with fruit availability, as does waiting time to conception (Emery Thompson, personal communication). In Kibale, conceptions are significantly associated with periods of high fruit availability whereas other reproductive states are not (Ellison et al. 2005).

Other primate species also show evidence of variation in ovarian function and fecundity associated with energy availability, but the evidence from chimpanzees and orangutans is particularly relevant to students of human evolution (Ellison et al. 2005; Knott 2001). It suggests that this pattern of variability is not a unique human characteristic and may well have been expressed in the last common ancestor of chimpanzees and humans. It also suggests that patterns of environmental energy availability may have been important selective factors in maintaining and organizing this aspect of female reproductive physiology.

HYPOTHESIS 3: EARLY FETAL LOSS MAY BE SENSITIVE TO ENERGY AVAILABILITY, BUT MAY ALSO BE SENSITIVE TO OTHER ASPECTS OF “STRESS” AS WELL

Reproduction in women is a process with several critical junctures. Ovulation is one such juncture. Without the production of a viable egg cell a pregnancy is not possible. With the production of an egg, fertilization may or may not occur. If fertilization occurs, implantation in the uterus may or may not take place. If implantation occurs, the pregnancy may or may not last long enough to be recognized as such by the woman or her physician. If a recognized pregnancy occurs, it may or may not result in a live birth. In a manner reminiscent of the continuum of ovarian function, we can think of this temporal sequence as having a “clinical horizon,” a point at which a pregnancy is recognized by its visible manifestations (Figure 4). This horizon is fuzzy, however, because it depends on the technology employed to detect a pregnancy. In a premodern context a missed menstrual period combined with physical symptoms like breast tenderness might have constituted the “clinical horizon,” the point at which a woman would recognize that she was pregnant. In contemporary clinical practice, a “clinical pregnancy” is recognized when a fetal heartbeat can be detected, either with an oscilloscope or ultrasound. But there are also ways to investigate the critical stages that precede
this clinical horizon. For example, as we have seen above, steroid profiles can be used to infer whether ovulation has taken place.

Some evidence indicates that an embryo begins to produce distinctive chemical signals even before it has implanted in her uterus, though reliably detecting these signals in the mother’s blood or urine is exceedingly difficult (Morton et al. 1992; Rolfe 1982; Rolfe et al. 1988; Shahani et al. 1992). Once an embryo implants in the endometrium of the uterus, however, it begins to secrete the hormone human chorionic gonadotropin (hCG) directly into the mother’s blood stream in increasing amounts. The gene for the distinctive beta subunit of chorionic gonadotropin (CG) is derived from that for the pituitary gonadotropin, luteinizing hormone (LH). The beta-CG gene first appeared in the common ancestor of anthropoid primates and was subsequently duplicated twice in the catarrhine primates (Maston and Ruvolo 2002). Among its other functions, LH stimulates progesterone production by the corpus luteum, which in turn maintains the endometrial lining of the uterus. HCG also stimulates hormone secretion by the corpus luteum, but even more effectively than LH and with a half-life in the blood stream that is nearly two orders of magnitude longer. It is crucial to the continuance of a pregnancy that the embryo send the hCG signal to reach the mother’s ovary almost as soon as it has established a connection to her blood stream, because without that signal the corpus luteum will begin regressing in a day or two and the resulting menstrual bleeding will carry away the embryo with it. Detecting hCG in the mother’s blood or urine therefore indicates the presence of an implanted embryo. Most home and hospital pregnancy detection tests are based on this determination. However, most clinics and even most home pregnancy kits will recommend that any “positive” test be repeated after several days or a week for the pregnancy to be confirmed. This is not because the tests themselves are invalid, but because the presence of hCG at an early stage does not yet indicate the “successful” establishment of a pregnancy. By comparing the evidence of early and regularly repeated hCG determinations with standard evidence of a “clinical pregnancy” later, a number of studies have indicated that the rate of embryonic loss in this interval is quite substantial (Edmonds et al. 1982; Wilcox et al. 1988). One of the best such studies, carried out in North Carolina by Wilcox et al. (1988), estimates that nearly one-third of those embryos whose presence can be detected early on from hCG measurements are lost before establishment of a clinical pregnancy.

Early embryonic loss of this kind must also be considered a possible contributing factor to variation in female fecundity (Macklon et al. 2002). We do not know much yet, however, about the degree to which early embryonic loss is associated with characteristics of the embryo or characteristics of the maternal environment, or both. Analysis of the

Figure 4. The temporal sequence of conditional stages leading to a live birth. Each stage is conditional on the success of early stages. The “clinical horizon” in this case is the point at which the existence of a pregnancy can be determined from detection of a fetal heartbeat.
North Carolina data set has indicated that steeper rises in progesterone metabolites in the urine in the week following implantation are associated with lower rates of early embryonic loss (Baird et al. 2003). There is clinical evidence from assisted reproduction studies that successful implantation (resulting in a clinical pregnancy) of developing embryos depends on adequate endometrial development (Dickey et al. 1993; Kovacs et al. 2003; Merce et al. 2007; Reuter et al. 1996; Zhang et al. 2005). Additional evidence indicates that endometrial preparation in turn may depend on the degree of estrogen stimulation in the follicular phase of the ovarian cycle, when the endometrium is proliferating (de Ziegler et al. 1998; Reuter et al. 1996). This is evidence for maternal factors having an effect. Notably, however, many of the results on endometrial factors that are significant in assisted reproduction, where hormonal levels are highly manipulated, are not found in studies on naturally occurring cycles (Check et al. 1995). But there is also clear evidence from in vitro fertilization results that embryos vary in “quality,” usually indexed by the rate of initial cell division, and that embryo quality is highly correlated with implantation success (Minaretzis et al. 1998). Embryo quality might be influenced by maternal factors such as follicular estrogen levels, and maternal age (Hourvitz et al. 2006; Shen et al. 2003), but almost certainly includes genetic factors endogenous to the embryo itself (Orvieto et al. 2004).

Whatever the source of variability in implantation success, evidence for patterns of variability that might indicate differential responsiveness to energetics is ambiguous. Holman (1996) found rates of early embryonic loss in Bangladesh that are nearly identical to those found in North Carolina using the same hCG assay methods. This evidence would suggest that there is little variation in early embryonic loss at the population level, but does not directly address variation within women, or between women within populations. Nepomnaschy and colleagues (2004; 2006), on the other hand, have reported that, for Guatemalan subjects, cortisol levels were significantly higher in women who experienced early embryonic loss than at comparable times in women in whom a clinical pregnancy was established. As noted above, cortisol is a catabolic hormone that functions to mobilize stored energy. It rises under conditions of energy shortage or energy demand. But it also can be elevated in response to psychosocial stress as part of a generalized stress response in anticipation of action that may require energy. Nepomnaschy et al.’s study suggests that there may be a connection between the catabolic mobilization of energy and early embryonic loss, at least between women. Hopefully additional information will add to this picture soon. At present, however, this hypothesis must be considered very tentative.

**HYPOTHESIS 4: FETAL GROWTH IS SENSITIVE TO ENERGY AVAILABILITY, ALTHOUGH HIGHLY BUFFERED**

There is considerable evidence that the birth weight of babies carried to term is sensitive to maternal energy avail-ability (Habicht et al. 1973; Kramer 1987). This is true at both the population and individual levels. Periods of famine and general nutritional deprivation are associated with reduced birth weight distributions (Lumey et al. 1995; Stein et al. 1995) while clinical and epidemiological evidence supports the association between energy availability and fetal weight gain within and between women (Frentzen et al. 1988; Rode et al. 2007; Shapiro et al. 2000). Third trimester fetal fat deposition is particularly sensitive to maternal energetics (Nyaruuhucha et al. 2006). In addition to the negative consequences of low maternal energy availability on fetal growth, negative consequences of very high energy availability in raising the risks of difficult labor and caesarean delivery have also been noted, even in developing countries (Onyiriuka 2006).

The relationship between maternal energetics and fetal weight gain is mediated by the dependence of the fetus on the glucose levels in the mother’s circulation for primary oxidizable substrates (Aldoretta and Hay 1995). Glucose, the most important of these substrates, passes from maternal to fetal circulation via facilitated transport, but at a rate that is proportional to maternal glucose levels (Morris and Boyd 1988). Conditions of elevated maternal glucose, such as gestational diabetes, are associated with accelerated fetal weight gain and higher birth weight (Ray et al. 2001). At the low end, however, maternal glucose levels may be buffered against falling too low even though maternal energy may be limited.

Compelling evidence of this comes from studies of the effect of caloric supplementation of undernourished Gambian women during pregnancy (Prentice et al. 1993; Prentice et al. 1987; Prentice et al. 1988; Prentice et al. 1983b). Mothers received energy supplements averaging 431 kcal/day during pregnancy, or an increase in average energy intake of more than 40% compared to a baseline period without supplementation. On average, birth weights increased only slightly, by about 50g, as a result. But birth weight to women who were pregnant during the energy-scarce wet-season improved more, 224g on average, and the rate of low birth weight babies (<2500g) in this group decreased from 28.2% to 4.7%. There was no significant effect of energy supplementation in women pregnant during the dry season, however, when women are normally in positive energy balance. It also was documented that women in this population frequently reduce their own basal metabolic rates when they become pregnant, freeing up metabolic energy for the pregnancy (Poppitt et al. 1993; Prentice et al. 1989; Prentice et al. 1995). Supplementation may simply have reduced the need for this trade-off. These observations suggest that maternal physiology works to buffer fetal growth under conditions of low energy availability in order to maintain a viable pregnancy, a pattern that has been observed in other study populations as well (King et al. 1994). This buffering may require trade-offs with other areas of the mother’s energy budget, however, as predicted by life history theory.
HYPOTHESIS 5: FAT STORAGE IN PREGNANCY IS A HIGH METABOLIC PRIORITY, ESTABLISHING RESERVES THAT CAN BE MOBILIZED IN LATE PREGNANCY AND DURING LACTATION

The Gambian data point up another important aspect of reproductive energetics in humans, the importance of fat storage in early pregnancy. Gambian women who are not supplemented when pregnant nevertheless gain weight (Prentice et al. 1981). This weight gain cannot be fully accounted for by increases in caloric intake or by decreases in energy expended in physical activity. Rather, it reflects the energy reallocated from their own basal metabolic budgets (Poppitt et al. 1993). The weight gained by a pregnant woman in the first trimester of pregnancy is primarily a reflection of fat storage, not placental or fetal tissue mass. The physiological response of Gambian women thus indicates that fat storage during this period is a very high metabolic priority, high enough to divert scarce energy streams to accomplishing it. This high priority probably is explained by the fact that the final weeks of pregnancy are typically characterized by rapid mobilization of maternal fat stores, even in women living under favorable energetic conditions (Homko et al. 1999; Sivan et al. 1999). Indeed, maternal metabolism in late pregnancy has been likened to a period of accelerated starvation as the mother’s body strives to keep up with the metabolic demand of her fetus (Freinkel 1980; Homko et al. 1999). It is likely that maternal metabolism shifts to fat utilization at the end of pregnancy in part to spare carbohydrate substrates for the fetus (Butte et al. 1999) as well as to meet the fetus’ lipid requirements (Neville 1999). Fat accumulated early in pregnancy may be instrumental in meeting this demand late in pregnancy and in lactation, particularly for women whose prepregnancy fat reserves are low to begin with. Villar et al. (1992) have confirmed the importance of maternal fat accumulation early in pregnancy to subsequent birth weight, finding that “Maternal nutritional status at the beginning of gestation and the rate of fat gain early in pregnancy are the two nutritional indicators most strongly associated with fetal growth.”

HYPOTHESIS 6: THE CASCADE OF SIGNALS THAT INITIATES LABOR BEGINS WITH THE MOTHER’S INABILITY TO FULLY MEET THE GROWING METABOLIC NEEDS OF THE INFANT

Birth does not mark the end of direct maternal investment in the offspring, but only a transition from investment in utero to investment ex utero. The optimal timing of this transition is sensitive to several factors related to energetics. The energy requirements of the fetus mount steeply at the end of pregnancy, driven both by the requirements of its rapidly growing brain and by the need to augment its own fat reserves to survive the neonatal period (Butte et al. 2004). Kuzawa (1998) and Dufour and Sauther (2002) have both pointed out that human babies are particularly fat at birth in comparison with other primates and that their fat reserves are important buffers for a comparatively altricial infant. But brain growth and fat deposition in the fetus at the end of pregnancy also present an increasing challenge to the mother, both because the total energy requirement of the fetus is so high and because there is a particularly high requirement for fatty acids, which do not cross the placenta well, to support both processes (Aldoretta and Hay 1995; Morris and Boyd 1988). As the mother’s ability to meet fetal metabolic requirements begins to fall behind the increase in those requirements, the fetus begins to release cortisol from its own adrenal gland and to mobilize, rather than build up, its own fat reserves. Cortisol release from the fetal adrenal gland is a sign that the end of the pregnancy is near and stimulates processes, like the secretion of surfactant in the fetal lungs that prepare the fetus for birth. It also has now been shown that fetal cortisol stimulates prostaglandin production in the placenta, liberating stores of arachidonic acid, an essential fatty acid required by the fetal brain, directly into the fetal circulation and initiating labor contractions in the uterus (Majzoub et al. 1999; McLean and Smith 2001; Smith et al. 2002). Thus the cascade of signals that leads to the initiation of labor begins with the inability of the mother to adequately meet the fetus’ metabolic requirements. This hypothesis has been termed the “metabolic cross-over hypothesis,” because it views the initiation of labor as a consequence of a cross-over in the curves reflecting fetal metabolic requirements and maternal ability to meet those requirements (Ellison 2001b, 2003) (Figure 5).

The transition to ex utero investment does not diminish those requirements, which in fact continue to grow as the infant does (Butte and King 2005). But the mother’s ability to meet those requirements is increased by the ability to pass lipids and fatty acids to her infant in her breast milk (Agostoni et al. 2001; Marangoni et al. 2002; Mitoulas et al. 2003; Ortiz-Olaya et al. 1996). Even though human breast milk is not as high in fat content as that of many other mammals, it does provide for a greater rate of energy transfer than can be accomplished across the placenta. The polyunsaturated fatty acids in the milk also are particularly important for infant brain development (Agostoni et al. 2001). The mother also is able to utilize her own fat reserves more efficiently to support milk production than she can to produce simple carbohydrates for placental transfer. Thus, although birth exposes the infant to numerous new challenges to its survival, it is a necessary transition to a new mode of nutritional support capable of meeting the infant’s increasing metabolic demands, especially those of its brain.

HYPOTHESIS 7: THE DURATION OF GESTATION IS SENSITIVE TO ENERGY AVAILABILITY, ALTHOUGH BUFFERED BY OFFSETTING EFFECTS

Hypothesis 6, the metabolic cross-over hypothesis, predicts that mothers with restricted energy availability will initiate labor at a lower threshold of fetal metabolic demand than women whose energy availability is not so restricted. This is consistent with Hypothesis 4, that more energy-restricted mothers will have lower birth weight babies than...
less restricted mothers. In general, more energy restricted mothers will reach the metabolic cross-over more quickly than less restricted mothers as well, resulting in shorter gestation lengths (Peacock 1991). This effect is somewhat buffered, however, by the fact that fetal growth is slower in energy restricted mothers, delaying the point at which any particular level of fetal metabolic demand is reached. These two effects do not fully offset each other, however, and there is substantial evidence that energy availability is positively correlated with gestation length. Clinically, syndromes like gestational diabetes that increase energy availability are associated with extended gestation (Cunningham et al. 1989). Epidemiologically, periods of energy restriction, like the “Dutch Hunger Winter” during the final phase of Nazi occupation in World War II, are associated with shorter gestation lengths of term pregnancies, discounting spontaneous abortions and miscarriage (Stein et al. 1975).

Note that this hypothesis is contrary to two commonly held, but erroneous, beliefs about gestation length. One is that parturition is determined by an “internal clock” that simply goes off at nine months. If undetected, the gestation of an anencephalic fetus (a fetus congenitally lacking a neocortex, and thus with particularly low metabolic requirements) can extend for more than a year (Higgins 1954). This indicates that there is nothing about the physiology of gestation that simply gives out as a matter of time alone. The second belief is that parturition occurs when the fetus reaches a certain state of maturity or “readiness.” If this were true, we would expect fast growing fetuses to be born earlier and slow growing ones later. In fact, just the opposite is true.
HYPOTHESIS 8: RESUMPTION OF FECUNDITY POSTPARTUM IS SENSITIVE TO ENERGY AVAILABILITY

Since the middle of the 20th century, it has been recognized that the period of lactation in humans is associated with a period of amenorrhea of variable duration (Ellison 1995; Henry 1961). A substantial part of the variance in the duration of lactational amenorrhea can be accounted for by variation in the temporal pattern of infant nursing (Howie and McNeilly 1982; Vitzthum 1989). In particular, Konner and Worthman (1980) argued that the long duration of lactational amenorrhea and associated long interbirth intervals observed among !Kung hunter-gatherers in Botswana could be explained by their pattern of frequent, on-demand breastfeeding that ordinarily continues for several years. In contrast, the pattern of low frequency breastfeeding practiced by many mothers in Western countries, they argued, was responsible for the relatively short periods of lactational amenorrhea they experienced. Data from an increasing number of traditional societies and developing countries, where mothers tended to nurse more frequently and experience longer periods of lactational amenorrhea than in the West, seemed to accord with this hypothesis, known as the “nursing frequency hypothesis” (Ellison 1995; Van Ginneken 1977). However, not all the data lined up satisfactorily. Variation in the duration of lactational amenorrhea between women within the same population, with similar nursing practices, was found to be substantial and often not explained by variation in nursing frequency (Brown et al. 1985; Huffman et al. 1987). The best predictor of the resumption of menstruation in carefully monitored, longitudinal studies in Scotland was the introduction of supplementary food into the infant’s diet, an event that was not well-correlated with changes in nursing frequency (Howie and McNeilly 1982; McNeilly et al. 1994; Tay et al. 1996). And comparison of populations with high nursing frequencies, even as high as the !Kung, revealed considerable variation in average duration of amenorrhea (Elias et al. 1986; Valeggia and Ellison 2004; Worthman et al. 1993).

Once again, an important insight can be derived from the Gambian supplementation studies. In addition to nutritionally supplementing women during pregnancy, some women also were given supplements during lactation in an attempt to increase rates of postnatal growth by increasing the amount and/or quality of breast milk produced by the mothers. Once again, the experiment largely failed in its intended objective, as neither the quantity nor the energy content of breast milk was affected by the supplementation (Prentice et al. 1983a). Unexpectedly, however, the supplemented women experienced significantly shorter periods of lactational amenorrhea and became pregnant again sooner than unsupplemented mothers. They also experienced a more rapid postpartum decline in levels of the hormone prolactin, a hormone that supports milk production by differentially partitioning metabolic energy to that end (Lunn et al. 1981; Lunn et al. 1984; Lunn et al. 1980). Based on these results, Lunn has proposed an alternative to the nursing frequency hypothesis, the “metabolic load hypothesis,” suggesting that it is the metabolic load on the mother’s energy budget that lactation imposes that is correlated with the suppression of ovarian function (Lunn 1992). The metabolic load can become absolutely lighter as an infant demands less milk, or it can be relatively lighter if the mother’s energy availability is higher. Nutritional supplementation in the Gambian mothers, he argues, effectively lowered the relative metabolic load of lactation and resulted in shorter periods of lactational amenorrhea.

A different intervention study recently has produced results that echo those of the Gambian supplementation studies. Gibson and Mace (2002, 2006) report on a program in Ethiopia aimed at making clean drinking water more readily available in rural locations by drilling local wells. The hope was that this intervention would lead to lower rates of child morbidity and mortality and faster growth rates. The wells did dramatically reduce the energy that women had to expend in hauling water from distant sources. But the effect was to shorten the period of lactational amenorrhea and the interval to the next pregnancy. Although infant and child mortality rates did decline, growth rates were actually slower where the wells were introduced. Gibson and Mace speculate that the larger family sizes that resulted from the combination of higher fertility rates and higher child survival rates actually reduced food intake per child within families, leading to slower growth.

Claudia Valeggia and I have explored the metabolic load hypothesis by following hormonal indicators of energy availability longitudinally in a sample of 70 nursing Toba mothers in Argentina (Valeggia and Ellison 2004; Valeggia and Ellison 2001). The Toba are an indigenous group living in the Gran Chaco area of northern Argentina. The Toba subsisted as hunter-gatherers until the early part of the 20th century, and now fall along a gradient of integration into the national market economy, largely depending on their proximity to large towns. The mothers we studied live in a government built settlement on the outskirts of the provincial capital of Formosa. In general, energy availability is very high in this population with a diet high in starch and fat and very low levels of physical energy expenditure (Valeggia and Ellison 2001). Adults and children are well-nourished by national and international standards, and average birth weight is high by the same standards (Faulkner et al. 2000). Toba mothers still nurse their offspring on demand, however, with frequencies that are the same as those reported by Konner and Worthman (1980) for the !Kung. Despite such frequent nursing, the average duration of lactational amenorrhea among the Toba is only 10 months compared to nearly two years among the !Kung (Valeggia and Ellison 2004).

The relatively short average duration of lactational amenorrhea among the Toba is in line with the metabolic load hypothesis. But we wondered whether variation in the duration of amenorrhea between Toba mothers also could be explained by differences in energy availability. We followed 70 mother-infant pairs from early in lactation until three months after the resumption of menstruation, observing patterns of breastfeeding, changes in weight, and
collecting frequent urine samples for hormonal analysis (Valeggia and Ellison 2001). In addition to measuring ovarian steroid levels in the urine samples, we also measured a molecule known as C-peptide of insulin, or simply “C-peptide” for short. C-peptide is a p-polypeptide chain that is cleaved from the proinsulin molecule to produce active insulin. Unlike insulin, it is cleared intact into the urine where it can be measured, allowing accurate inferences about insulin production. Diana Sherry and I have pointed out that C-peptide can be interpreted as a marker of energy availability, since elevated insulin promotes energy storage and anabolic allocation and opposes catabolic processes, while suppressed insulin has the opposite effect (Sherry and Ellison 2007).

Among Toba women in lactational amenorrhea C-peptide levels are initially low compared to levels observed in the same women when they are menstruating (Figure 6). But as the postpartum period progresses, C-peptide levels steadily rise (Ellison and Valeggia 2003). In fact, in the one to two months immediately prior to the resumption of menstruation, C-peptide levels are significantly higher on average than they are after menstruation resumes. There is a high temporal correlation between the peaking of individual C-peptide trajectories and the resumption of menstruation, and there is also a close correlation between increasing C-peptide levels, increasing maternal weight, and increasing urinary estrogen levels during the months preceding menstrual resumption. On the other hand, there are no correlations between C-peptide levels and any indices of nursing pattern or frequency (Figure 7). We interpret these data as indicating that variation in the duration of lactational amenorrhea within the Toba population is sensitive to differences in energy availability within and between women. Indeed, we have even suggested that increasing insulin levels may be causally connected to the stimulation of steroid production in the ovary (Franks et al. 1999; Poretsky and Kalin 1987).

It should be stressed that the nursing frequency hypothesis and the metabolic load hypothesis are not in essential opposition to one another. Rather, the metabolic load hypothesis encompasses the nursing frequency hypotheses to the extent that variation in nursing frequency can often be associated with variation in the metabolic load of lactation. The empirical problem is that such an association can result either from a direct effect—a baby that nurses more frequently may actually consume more milk than one that nurses less frequently—or from a confounding of nursing frequency with maternal energy availability—women in traditional societies may both nurse their babies more frequently and have lower energy availability than their peers in Western societies. But while the metabolic load hypothesis encompasses the nursing frequency hypothesis, the reverse is not true. The nursing frequency hypothesis cannot explain why populations or individuals

![Figure 6. Average (± SE) values of urinary C-peptide from 70 Toba women, expressed as a percentage of each woman’s individual average value when menstruating and aligned on the month of menstrual resumption (Ellison and Valeggia 2003).](image)
with the same nursing pattern can have very different durations of lactational amenorrhea.

**HYPOTHESIS 9: MALE FECUNDITY IS INSENSITIVE TO ENERGY AVAILABILITY, AT LEAST IN THE NORMAL RANGE, BUT OTHER ASPECTS OF MALE REPRODUCTIVE EFFORT MAY BE SENSITIVE TO ENERGY AVAILABILITY**

All of the hypotheses reviewed to this point concern the relationship of energetics to female reproduction. In part this is because more research has been conducted on female than on male reproductive ecology, in part because female fecundity has a greater impact on population level fertility than does male fecundity, and in part because the available evidence indicates that female reproductive effort is more sensitive to energy availability than is male reproductive effort. The latter two reasons are no doubt partially responsible for the first. The latter two reasons also are predictable from evolutionary theory. Because female mammals, including humans, are physiologically saddled with bearing the direct metabolic cost of gestation and lactation, energy availability becomes an important limiting factor for their reproductive success. The fact that males, in contrast, can simultaneously father many offspring means that male reproductive success is less limited by energy availability and more limited by mating access to females.

Male fecundity, if we control for frequency of intercourse as we do for female fecundity, is primarily a function of sperm production, both quantity and quality (Lamb and Bennett 1994). It appears that both of these aspects of male reproductive physiology are relatively insensitive to energy availability within a very broad range (Bribiescas 2006). Even marathon training has little appreciable effect on sperm counts (Bagatell and Bremner 1990). The probability of conception per act of intercourse in the woman’s fertile period is relatively insensitive to sperm count in any case, unless it falls to abnormally low levels (Wood 1994).

But while male fecundity may not be sensitive to energy availability, other aspects of male reproductive effort may be. Bribiescas (1996, 2001, 2006) has noted that testosterone levels help to modulate muscle mass, which can be considered a form of somatic reproductive effort in males, con-

---

**Figure 7.** Average normalized C-peptide values for 70 Toba women compared with average values for (A) body mass index (kg/m²); (B) urinary estrone conjugates (ng/mg creatinine); (C) nursing bout frequency (bouts/hr); and interval between nursing bouts (mins).

tributing to success in male-male competition for resources and/or access to females. Although male testosterone levels do not fluctuate appreciably in response to short-term or low amplitude fluctuations in energy availability (Bentley et al. 1993), there is evidence that they are sensitive to longer term and higher amplitude variation (Ellison et al. 1989; Klibanski et al. 1981). This may be one of the mechanisms by which males modulate their reproductive effort to fit their overall energy budget. There is also evidence that this somatic form of reproductive effort is linked to trade-offs with metabolic investment in survival and maintenance functions, such as immune function (Muehlenbein and Bribiescas 2005). Testosterone also may be related to behavioral aspects of reproductive effort in males, including support of libido and the pursuit of dominance and social status (Bribiescas 2006; Dabbs 2000; Ellison 2001b; Mazur and Booth 1998). Recent studies even indicate that in situations where men invest heavily in mates and offspring, the shift from mating effort to parenting effort is associated with a shift to lower testosterone levels (Gray et al. 2002; 2004).

Thus, testosterone may have an important role in modulating male reproductive effort and the energy it involves even if it has little impact on male fecundity per se. These other aspects of male reproductive effort may be much more important in terms of the energy they require than is sperm production, and to that extent they may show greater sensitivity to energy availability.

HYPOTHESIS 10: AGE AT REPRODUCTIVE MATURATION IS SENSITIVE TO ENERGY AVAILABILITY, BUT AGE AT REPRODUCTIVE SENESCENCE IS NOT
Although age at reproductive maturity, like adult height, is highly heritable so that early maturing parents tend to have early maturing offspring, it is also very sensitive to environmental factors, and in particular to energy availability (Eveleth and Tanner 1991; Norgan 2002). This is true for both sexes, although there is more abundant data for females, largely because age at first menstruation—menarche—is such an easily collectible datum. In both males and females, age at the adolescent growth spurt is highly correlated with age at reproductive maturation because both are consequences of gonadal steroid production (Ellison 2002; Hauspie 2002; Hindmarsh 2002). Any of these markers of reproductive maturation show clear patterns of variation with indices of energy availability such as tempo of childhood growth or body mass index, and variation with other factors, such as socioeconomic status, that correlate with energy availability (Johnston 2002). Secular acceleration of maturation is widely observed as energy availability increases within populations over time, or as a consequence of migration from low to high energy availability circumstances (Bogin 2001; Eveleth and Tanner 1991). A reasonable summary statement is that, to the extent that energy availability influences childhood growth, it also influences age at reproductive maturation.

Although the onset of the reproductive phase of life is sensitive to energy availability, its termination is not, or is much less so. Ovarian function does begin to decline in the late thirties, but the trajectory of decline appears to be similar across a range of different populations and ecologies (Ellison 1996a, b) (Figure 8). Menopause itself is a consequence of follicular depletion in the ovaries (Vom Saal and Finch 1988). Depletion occurs because the supply of follicles is finite, having been determined in utero, while the rate of follicular atresia is relatively constant from midgestation to the menopausal transition, and independent of either energy availability or any aspect of ovarian function. Follicular depletion is an unavoidable occurrence in any mammal that lives long enough, and is well documented in captive primates (Brenner et al. 2004; Nichols et al. 2005). Follicular depletion occurs at about the same time in chimpanzees as in humans, perhaps related to the similar size of the ovaries in the two species (Jones et al. 2007). The unusual feature of human life histories related to menopause is not that it occurs or when it occurs, but that humans routinely live well beyond it.

In men, age-related decline in free testosterone levels begins as early as the mid-twenties and is appreciable by the mid-forties. As with levels of ovarian steroids in women, the age-specific mean testosterone values for different populations can vary, particularly in young adulthood. But by age 60 the average levels across different populations tend to converge (Ellison et al. 2002) (Figure 9). The rate of change in testosterone with age is related to body composition change, with relative decreases in muscle mass and increases in fat mass, reflecting a shift in somatic energy allocation away from somatic reproductive effort with increasing age (Campbell et al. 2006; Lukas et al. 2004; Vermeulen et al. 1999). This rate of decline may be sensitive to energy availability, but in an unexpected direction, because it appears that men in populations characterized by high energy availability have steeper rates of decline in testosterone and change in body composition than do men in populations with lower energy availability (Ellison et al. 2002). This fact is less mysterious, however, if we consider that energy availability has a stronger influence on testosterone levels in young adulthood and much less of an influence late in life.

HYPOTHESIS 11: ENERGY AVAILABILITY EARLY IN DEVELOPMENT MAY INFLUENCE REPRODUCTIVE EFFORT IN ADULTHOOD
In a review of working hypotheses in the newly emerging field of human reproductive ecology, similar to this one, published in 1990, I included the following hypothesis as number 8: “Late reproductive maturation is associated with a slower rise in indices of ovarian function with age, and a lower level of ovarian function in adulthood” (Ellison 1990). Later I noted, “The most likely functional interpretation of hypothesis 8 in this context is that slow physical development and reproductive maturation provide a kind of ‘bioassay’ of chronic qualities of the environment, particularly energy availability” (p. 944). I also identified this hypothesis as “perhaps the most controversial of the set”
Seventeen years later all three statements continue to be relevant. The data pertaining to them have increased, however, allowing the hypothesis to be expanded somewhat in scope to include males as well as females and to encompass in utero as well as postnatal developmental effects. Controversy continues nevertheless, both over the empirical nature of these effects and over their functional and theoretical interpretation.

New interest in early developmental effects on adult biology has been generated by the “Barker hypothesis,” also known as the “fetal origins hypothesis.” Barker and his colleagues drew attention to the fact that the risk of many chronic diseases that appear in adulthood is correlated with size at birth, and suggested that low energy availability during gestation might force developmental trade-offs in utero that would have functional consequences for adult health (Barker 1994). Gluckman and Hanson (2005) have argued that these sorts of developmental linkages might have adaptive significance. Terming their hypothesis the “predictive adaptive response” (PAR) hypothesis, they suggest that conditions in utero might provide important cues to the expected postnatal environment and that the developmental trajectory of the organism may shift as a consequence in an effort to align its physiology with these expected conditions (Gluckman et al. 2005). In particular, they argue that high energy availability in utero, leading to a larger birth size, results in a shift to a trajectory of faster growth, earlier maturation and a less “conservative” adult metabolism, while low energy availability in utero has the opposite effect. When there is a “mismatch” between conditions in utero and after birth, disease outcomes may result. In particular, when low energy availability in utero is followed by high energy availability after birth, a tendency to obesity and the metabolic disease syndrome can result (Gluckman and Hanson 2004).

Empirical data on in utero effects on later development are complex and not always consistent. The effects of small size at birth are often confounded with the effects of rapid, or “catch-up,” growth in infancy, but when data on both are available they tend to support the PAR hypothesis (Adair 2001; Ekelund et al. 2007; Ong 2006). The effects of birth size on reproductive physiology are still sparse and not always consistent. Small birth size has been linked to reduced uterine and ovarian size, but also to early age at menarche (Adair 2001; Ibañez et al. 1998; 2000; 2002). The latter association may be an example of confounding with rapid postnatal growth, however, and not observed in small babies who continue to grow slowly (Ong 2006).

Jasienska et al. (2006b) have presented the only data so far linking size at birth to adult ovarian function, finding that Polish women who were larger at birth have higher average estradiol levels in adulthood than peers who were
smaller at birth. They also have reported that size at birth interacts with energy expenditure in adulthood in influencing ovarian function, with estrogen levels in women who were large at birth apparently being less sensitive to energy expenditure (Jasienska et al. 2006a) (Figure 10). They note that this evidence of a shift in the norm of reaction of ovarian function to energy availability in conjunction with birth size provides an affirmative test of the PAR hypothesis.

Not all new evidence points to the importance of in utero effects on adult ovarian function, however. Nuñez de la Mora et al. (2007) recently reported on a study of ovarian function in Bangladeshi migrants to London compared to women remaining in Bangladesh and London women not of Bangladeshi extraction. As might be expected from research on other populations, Bangladeshi women in Bangladesh have profiles of salivary progesterone that are significantly lower than those of London women. The daughters of Bangladeshi migrants to London who were born in London have progesterone profiles that are indistinguishable from those of other Londoners, while Bangladeshi women who migrated to London as adults have profiles that are similar to their peers who remain in Bangladesh. Bangladeshi women who migrate to London as children, however, have levels that are intermediate between these extremes. When this group is broken down further, women who migrate before the age of nine have significantly higher progesterone profiles than those who migrate after age nine, the latter being comparable to women who migrate as adults. These results suggest the importance of environmental effects experienced in childhood, as opposed to before birth or after maturity, on adult ovarian function. The study does not identify, however, what aspects of the environmental difference between Bangladesh and London is implicated as a potential cause of the observed effect, though the authors speculate on the possible significance of different rates of parasitism and other forms of disease in the two environments.

Comparable data on males are not yet available. However, Bribiescas and others (Bribiescas 1996, 2001; Lukas et al. 2004) have argued that chronic energy availability during growth and development in part determines adult testosterone levels in men. This hypothesis is consistent with the population differences in testosterone levels in young adulthood noted above. More information on this question

Figure 9. Comparison of age trends in morning testosterone values for men from four populations, Boston men, Lese men from the Ituri Forest of the Democratic Republic of the Congo, Tamang men from Nepal, and Ache men from Paraguay. Solid lines represent the results of simple linear regression of morning testosterone against age for each population. For details of data and analysis, see Ellison et al. 2002.
is needed, however.

The functional significance of developmental effects on reproductive function remains a matter of debate. Kuzawa (2005) has questioned whether the intrauterine environment provides a reliable enough signal of postnatal conditions to be a basis for setting developmental trajectories. Constrained development in utero simply may have unavoidable consequences that are not necessarily adaptive. More efforts to discriminate these alternatives, comparable to Jasienska et al. 2006a, are needed. Vitzthum (2001; Vitzthum et al. 2004) has expressed doubt that developmental changes in adult ovarian function are functional, suggesting that average fecundity in different populations may be relatively similar despite differences in average levels of ovarian function and that it is only variation about the average level within populations that has functional significance. The data to test this alternative are lacking, as noted, pending good studies of fecundity across populations.

The hypothesis that energy availability has developmental effects on reproductive effort must therefore be regarded as provisional. It promises, however, to be an area of exciting research in coming years.

**REPRODUCTIVE ECOLOGY AND HUMAN EVOLUTION**

The working hypotheses just reviewed form a set of first-order conclusions that have emerged from the results of many empirical studies. Together they provide a context for further theorizing about the role of energetics in shaping our evolutionary history (Ellison 2001a).

One interesting characteristic of human reproduction seems to be the importance of energetic conditions around the time of conception. Ovarian function responds to changes in energy availability on a relatively short time scale on the order of individual ovarian cycles (Figure 11). As a result, when energetic conditions in the environment change seasonally, as in many subsistence agricultural societies, conceptions become seasonal as well with a nadir that is temporally correlated with the nadir in energy availability (Bailey et al. 1992; Ellison et al. 2005) (Figure 12). This form of reproductive seasonality differs significantly from the reproductive seasonality of many species in middle and high latitude environments. Seasonally breeding birds and mammals in these environments usually synchronize their reproduction in ways that synchronize birth or hatching and the immediately following period with the season of greatest energy availability (Bronson 1989). Migratory birds will even fly thousands of miles in order to synchronize their brood hatching with the brief but prodigious food abundance of the arctic tundra. Mating and fertilization may occur soon after arrival on the breeding grounds when females are near the nadir of their own energetic status. Early observers of human birth seasonality often tried to understand the phenomenon in these terms, hy-
Pothesizing that humans had developed cultural systems to increase the likelihood of birth occurring during seasons of abundance (see review in Ellison et al. 2005). The data indicate that this is not the case, however. It is the probability of conception, rather than the probability of birth, that shows a positive correlation with energy availability. Birth seasonality, per se, is an epiphenomenon of conceptions tending to be seasonally synchronized in environments with strong seasonality of resource abundance.

As noted above, this sensitivity of conception probability to energy availability is not unique to humans. It seems at the least to be shared by chimpanzees and orangutans among other hominoids (Ellison et al. 2005; Knott 2001). Why do we and our close hominoid relatives depart from the more familiar pattern of seasonally synchronized birth or hatching? Seasonally synchronized birth evolves when there is a concentrated, critical period in energy demand for successful reproduction and a predictable peak in environmental energy availability. Then, in the absence of other constraints, natural selection can be expected to favor individuals who can synchronize these two peaks, timing their conceptions so as to have the critical period of peak demand (often the early period after birth or hatching) occur during the period of peak abundance. Neither of these necessary preconditions may be true for humans and other hominoids. The profile of high energy demand tends to be broad and attenuated across gestation and lactation rather than concentrated in a brief, discrete period (Lee 1999; Prentice and Whitehead 1987). And, the environments in which humans and other hominoids have evolved, although seasonal to varying degrees, may not have been predictable.

But the absence of conditions that would select for seasonality of birth does not explain why the probability of conception should be sensitive to energy availability. If environmental energy availability is unpredictable, and the energy requirements of reproduction are spread over a long rather than a short time, why should conception be energy dependent at all? Again, a number of factors may contribute to an answer.

The simplest answer is that, unlike conditions in the distant future, energetic conditions in the present are known and probably predict the near future. Given that reproduction is energetically expensive, it may be better to synchronize it with periods of energy abundance to the extent possible rather than ignore energy availability altogether. “Catching a wave” of energy availability may make reproductive success more likely.

However, no matter how great the abundance of energy, if a period of abundance is short, the value of conceiving then may not be very great unless the energy harvested can be stored and mobilized later. At the start of a pregnancy in particular the immediate increase in energy requirements is very low, but later in the pregnancy it will be many times greater. The Gambian data underscore just how important the ability to accumulate stored energy early in pregnancy is for humans—important enough that women whose own energy budgets are marginal to begin with will downregulate energy allocation to their own maintenance and survival needs even further in order to free up energy for storage.

Fat storage provides the greatest energy depot in humans and other animals, with far greater capacity than glycogen and negligible maintenance costs. The principal value of fat storage is the ability to “smooth out” fluctua-
tions in energy availability by taking advantage of periods of abundance to subsidize periods of dearth. Other animals that regularly face long periods of starvation, including periods of hibernation, develop substantial fat reserves as buffers. Humans also have a considerable capacity to survive periods of famine based on their ability to mobilize fat. Brain tissue, which cannot directly metabolize fatty acids, can metabolize ketone bodies, by-products of fat catabolism. As a result, both brain and soma can be sustained for periods of weeks or even months by the slow mobilization of fat reserves. Compared to most other primates, humans have substantially higher fat reserves under “normal” (i.e., non-obese, non-emaciated) conditions (Pond and Mattacks 1987). And, although men carry less fat as a percentage of their total body weight than women do, the absolute amount of fat the average man carries is not that much less than the amount the average woman carries in most populations. (The differential in percentage of fat is more a consequence of large absolute differences in muscle mass between the sexes than in large absolute differences in fat mass.) This fact suggests that the value of fat storage as a buffer against variation in environmental energy availability has been particularly great in human evolution and may distinguish our formative ecology from that of chimpanzees, our closest living relative.

But in addition to using fat storage to smooth out fluctuations in environmental energy availability, human females use it to smooth out the fluctuations in energy demand that are imposed by pregnancy and lactation. Other animals typically increase energy consumption to meet the costs of pregnancy and lactation (Prentice and Whitehead 1987). Women in traditional societies, however, neither increase energy intake nor decrease energy expenditure enough to subsidize the costs of reproduction on a day-to-day basis (Dufour and Sauther 2002). They also rely on the ability to store fat when the energy demands of reproduction are low and to mobilize it when demand is high (Prentice et al. 1994). The lowest period of energy demand for reproduction is before conception even occurs and it remains nearly as low during the first weeks and months of pregnancy. Hence the value of having female fecundity be sensitive to energy availability to increase the chances of “catching a wave” of energy availability at the start of a reproductive bout.

The value of fat storage thus lies in the ability to smooth out fluctuation in net energy availability, whether the magnitude of this net value depends on variation in energy intake or variation in energy expenditure. It is likely that the enhanced (compared to chimpanzees) human capacity for fat storage evolved originally to smooth out fluctuations in

Figure 12. Percent of conceptions by month over ten years among Lese women (Bailey et al. 1989).
intake, since it is a capacity shared by both sexes. In human females, however, this physiological capacity appears to have been exapted to smooth out the fluctuations in energy expenditure imposed by pregnancy and lactation.

The value of fat storage and mobilization to smooth out either fluctuations in energy intake or fluctuations in energy expenditure should vary depending on the amplitude of those fluctuations and the net deficits in energy that would be faced without such mechanisms. The increased fatness of adult humans of both sexes relative to chimpanzees suggests that formative human environments involved greater amplitude fluctuations in energy availability than do chimpanzee environments. This is an inference that fits well with the resource implications of a shift from a rainforest to a savanna habitat and from a highly frugivorous diet with relatively abundant herbaceous fallback foods to a diet with increasing reliance on patchily distributed meat and tuber resources. If this inference is correct, then we also might infer that human ancestors increased their reliance on fat storage and mobilization during this habitat and dietary transition. Although orangutans are more distantly related to us than chimpanzees, their ecology, with its high amplitude variability in energy availability, may produce some of the same selection pressures favoring fat storage and mobilization as encountered by our hominid ancestors. A better understanding of their reproductive ecology, therefore, may well bring important insights.

As has been frequently noted, increases in brain size in the human lineage imply increases in the metabolic requirements for brain maintenance (Leonard et al. 2003; 2007). Less often noted is the fact that the energy requirements for brain growth also increase with brain size, with the bulk of that energetic burden falling on mothers. We saw evidence above that the energy requirements of fetal brain growth actually become limiting on the continuance of gestation at the end of human pregnancy, leading to the birth of a neonate that is markedly altricial compared to a chimpanzee neonate. The brain continues to grow rapidly during the first two years of postnatal life, with much of that growth continuing to be directly subsidized by maternal physiology through lactation. A reasonable inference from this evidence is that the amplitude of the fluctuation in energy demand imposed by gestation and lactation, relative to the overall maternal energy budget, also increased with the expansion of brain size during human evolution. The capacity for fat storage and mobilization as part of the female reproductive physiology would have become even more important and the sensitivity of female fecundity to energy availability even more valuable. Not only the ability to store fat early in pregnancy, but also the amount that a female is able to store during this period may be critical to the probability of reproductive success.

Enhanced capacity for fat storage and mobilization may thus have originally been selected for as a physiological mechanism for smoothing out high amplitude variability in environmental energy availability when early hominids shifted habitat and diet. It may have become further incorporated into female reproductive physiology to smooth out the increased amplitude variation in energy demands made by gestating an offspring with an increased brain size. This scenario differs from, and adds to, other hypotheses that also emphasize that the brain is “expensive tissue” by stressing the energy requirements of the growing brain, compounding what is already an energetically demanding state for a female. Reproductive ecology tends to stress the dynamic management of energy and energy allocation trade-offs that happen in dynamic, temporal relationship to one another. This perspective contrasts with the atemporal depiction of energy trade-offs presented by standard “expensive tissue” hypotheses. The relative cost of gut or muscle tissue versus brain in these hypotheses is examined as a general trade-off without reference to any specific developmental phase other than adulthood or to any temporal pattern of energy availability. One can summarize this contrast by saying that reproductive ecology presents a theory of energy dynamics while the standard “expensive tissue” hypotheses are based on a theory of energy statics.

Other aspects of human reproductive ecology reviewed above also stress energy dynamics and the allocation of energy to reproductive effort over time, especially over the course of development and through different life history stages. It emphasizes that reproductive effort is variable and facultative with patterns of change that map onto temporal dynamics of energy availability in the environment and developmental dynamics of growth, maintenance, and reproductive state. Particularly intriguing at the moment are the studies that suggest long-term and transgenerational developmental effects. Incorporating such a dynamic view of reproductive effort and energetics in general into our thinking about human evolution can inform both the scenarios we postulate for the past and our understanding of human biology in the present.

ACKNOWLEDGEMENTS

I would like to thank all the women and men of Argentina, Boston, Congo, Nepal, Paraguay, and Poland who participated in the studies described here. I am indebted to many colleagues for their collaboration, including Emily Barrett, Gillian Bentley, Rick Bribiescas, Ben Campbell, Judith Flynn Chapman, Peter Gray, Carol Hooven, Grazyna Jasienska, Sonya Kahlenberg, Cheryl Knott, Susan Lipson, Matthew McIntyre, Martin Muller, Mary O’Rourke, Catherine Panter-Brick, Meredith Reiches, Diana Sherry, Melissa Emery Thompson, Inger Thune, Claudia Valeggia, and Richard Wrangham. Financial support for the studies described here has come from NIH, NSF, DOD, the National Geographic Society, the Wenner-Gren Foundation, the Norwegian Cancer Society, and the John S. Guggenheim Foundation.
REFERENCES


Holman, D.J. (1996) *Total fecundity and fetal loss in rural Bangladesh*. University Microfilms, Ann Arbor, MI.


Nyaruhucha, C.N., Msuya, J.M., Ngowi, B., and Gimbi,


