Endocrine Variation in a Population of Peromyscus leucopus with Divergent Life-History Strategies

Jordan White
College of William and Mary

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

Recommended Citation
https://scholarworks.wm.edu/honorstheses/356

This Honors Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
Endocrine Variation in a Population of *Peromyscus leucopus* with Divergent Life-History Strategies

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Biology from The College of William and Mary

by

Jordan Thomas White

Accepted for Honors

(Honors)

Paul Heideman, Director

Eric Bradley

John Griffin

Randolph Coleman

Williamsburg, VA
April 25, 2011
**Abstract:**
The study of life-history strategies asks how the timing of reproduction, fecundity, offspring size, and related measures affect fitness. This thesis attempts to answer the following question: Does a change in hormone levels correlate with or cause a change in life-history strategies of an organism? Using two selection lines of *Peromyscus leucopus*, we investigated the variation in several hormone levels between the two mouse lines. The responsive line responds to a short day length by suppressing reproduction. The non-responsive line maintains reproduction in short days. There were no significant differences between lines in their mean levels of peripheral Insulin-like Growth Factor-1 (IGF-1). However, levels of IGF-1 were correlated with a number of reproductive measures in our mice. We conclude that levels of IGF-1 hormone, while related to reproduction, are not a likely source of evolutionary change in our mice. There were significant differences in levels of testosterone between photoperiods, but not between lines. We conclude that levels of testosterone are related to the photoperiod change, but not the difference between our selection lines. There were no significant differences in levels of estrogen between photoperiods or lines. However, serum effects may have affected the estrogen assay that was used. We conclude that there may be no significant differences in levels of estrogen, but these results should be confirmed with a more reliable assay.

**Introduction:**
Genetic variation in the control of reproductive hormones is a potential cause of variation in life-history strategies (Hau 2007; Williams 2008). Studies in birds have shown that hormone concentrations can correlate with specific life-history strategies (Hau et al. 2010). However, there is a dearth of mammalian data connecting heritable variation in hormone levels to heritable variation in life-history strategies (Hau 2007; Heideman et al. 2010). In order to test whether hormone levels might be related to evolution of life-history strategies we decided to test two hypotheses: the evolutionary constraint hypothesis and the evolutionary potential hypothesis (Hau 2007).

The evolutionary constraint hypothesis states that hormone receptor systems vary little and that genetic variation in levels of hormones is a more likely source of evolutionary change (Hau 2007). The evolutionary potential hypothesis states that hormone systems have numerous elements, such as receptors and binding proteins, which may evolve independently of each other. In the case of the evolutionary potential hypothesis, evolutionary change in phenotype could occur through many different genotypic and physiological changes, including evolutionary change in hormone levels. In order to test these hypotheses, we used a population of a temperate zone rodent that contains heritable variation in response to short day length.

For many small mammals, reproductive inhibition is common in the short days of winter. The harsh conditions of winter make winter reproduction risky for some males and females because of the added thermoregulatory costs (Fournier et al. 1999; Prendergast et al. 2001; Hill et al. 2008). However, within some
populations of *Peromyscus leucopus*, heritable variation in the photoperiod pathway can lead to reproduction in the winter (Heideman et al. 1999). Our laboratory maintains two lines of *Peromyscus leucopus* from such a population.

The two lines were artificially selected for either reproductive inhibition or non-inhibition in a short day photoperiod (Heideman et al. 1999). One line is Responsive (R) to photoperiod and suppresses its reproduction in a winter Short Day (SD). The other line is Non-Responsive (NR) to a short day photoperiod and maintains reproductive ability in short days. NR males and females have larger gonads than R males and females, respectively, in a SD photoperiod (Heideman et al. 1999; Broussard et al. 2009).

The physiological basis for the different phenotypes observed in NR and R mice could occur at numerous points in the complicated photoneuroendocrine pathway (Ebling and Barrett 2008). Neurons in the retina that perceive light send signals through a complex pathway to the pineal gland (Bartness et al. 1993; Goldman 2001). The pineal gland releases melatonin only during the night, providing a signal that allows the body to determine the duration of night and day (Bartness et al. 1993; Goldman 2001). One of the targets of melatonin is a group of hypothalamic neurons that release a reproductive hormone, gonadotropin-releasing hormone (GnRH) (Hadley and Levine 2007). GnRH causes the anterior pituitary to release luteinizing hormone (LH) and follicle-stimulating hormone (FSH). LH and FSH regulate gonadal development, with LH regulating the production and release of the sex-steroids, estrogen and testosterone (see figure 1 for a model). Sex steroids prepare an organism for reproduction and cause negative feedback regulation of GnRH neurons (Hadley and Levine 2007).

---

**Figure 1:** A diagram of the gonadotropin hormone axis in males. The axis is similar in females, except that in females these hormones are highly variable over time and in females estrogen plays a greater role than testosterone. GnRH = gonadotropin releasing hormone, FSH = follicle stimulating hormone, LH = luteinizing hormone, T = testosterone. Arrows with arrowheads are stimulatory. Arrows ending in bars are inhibitory.
A previous study has shown that non-responsive, female mice have higher levels of LH than responsive females (Heideman et al. 2010). Therefore, it is reasonable to predict that there is a difference in estrogen levels of females between the lines (for the broader concept see (Adkins-Regan 2008)). Similarly, one might expect a difference to occur between testosterone levels of NR and R males. A test of these predictions was one of goal of this study.

If these predictions are correct (estrogen and testosterone vary with respect to reproductive phenotype) then it would provide evidence for the ‘evolutionary constraint hypothesis’ (Hau 2007). However, if we find no differences between the steroid hormone levels in our selection lines, then we would have evidence against the evolutionary constraint hypothesis and for the ‘evolutionary potential hypothesis’ (Hau 2007).

A second goal of this study was to investigate a possible link between food intake differences in our lines and reproductive competence. When developing the selection lines, mice were selected only based on gonad size and development in short days. However, an apparently unrelated trait, food intake (but not body mass), of NR mice was found to be greater than food intake of R mice (Heideman et al. 2005). In order to investigate possible causes of the observed differences, we quantified levels of a hormone related to food intake and reproduction, insulin-like growth factor-1 (Daftary and Gore 2005; Veldhuis et al. 2006).

Insulin-like Growth Factor-1 (IGF-1) is a peptide hormone released peripherally by the liver (see figure 2). This peripheral IGF-1 is able to cross the blood brain barrier and to localize to the forebrain most abundantly (Reinhardt and Bondy 1994). Centrally, IGF-1 stimulates the release and production of Gonadotropin Releasing Hormone (GnRH) (Daftary and Gore 2005), which is a potent stimulator of the reproductive system. In addition to reproductive cycles, IGF-1 is also a known regulator of maturation and growth (Hoeflich et al. 2004; Veldhuis et al. 2006; Todd et al. 2007; Villalpando et al. 2008).

Figure 2: A diagram of the growth hormone axis. GHRH = growth hormone releasing hormone, GH = growth hormone, IGF-1 = insulin-like growth factor-1, GnRH = gonadotropin releasing hormone. Refer to figure 1 for information on GnRH activity.
Our null hypothesis was related to the evolutionary potential hypothesis: levels of IGF-1 hormone are not likely to evolve with a change in phenotype. Our prediction was that we would measure no difference between IGF-1 levels of the two mouse lines. We used the evolutionary constraint hypothesis as our alternative hypothesis: suggesting that IGF-1 levels can evolve to cause a change in reproductive phenotype. Our prediction was that the R line of mice would have lower concentrations of IGF-1 in SD relative to the NR line. This prediction follows from the results of a study by Ross et al (2009) in which a decrease in IGF-1 occurred in photoresponsive rats in SD photoperiod relative to LD photoperiod (Ross et al. 2009).

The prediction of our alternative hypothesis could prove false if the IGF-1 system has a low evolvability (McGlothlin and Ketterson 2008). Additionally, other hormone systems could cause the observed phenotypic differences between the mouse lines. In the latter case, we would predict no difference in IGF-1 levels between the mouse lines.

Methods:

Animals and Care:

The population of P. leucopus used was captured from a wild population in Williamsburg, VA in 1995 (latitude 37.3°N, longitude 76.7°W). The population was subsequently divided into subpopulations by selection across multiple generations (Heideman et al. 1999). Mice were selected for either reproductive inhibition in a Short Day photoperiod or the lack of reproductive inhibition in SD (Heideman et al. 1999). After being raised in a SD photoperiod, mice were examined at 70±3 days of age and assigned a reproductive index based on testis size or the size of the ovaries, uterine diameter, and presence or absence of visible corpora lutea (Heideman et al. 1999). Females with ovaries ≤2 mm in length, lacking visible corpora lutea, and uterine diameter of ≤0.5 mm were classified as reproductively inhibited (R) by SD. Females with large ovaries (usually >3.5 mm in length), large visible follicles or corpora lutea, and uterine diameter >1 mm were classified as nonresponsive (NR) to SD. Males with a testis index (length*width of testis) <24 mm² were classified as R; those with a testis index >32 mm² were classified as NR (Heideman et al. 1999). R males and females were paired in LD to produce offspring for the R line. NR males and females were paired to produce offspring for the NR line (Heideman et al. 1999).

Experimental Design:

IGF-1: Sample size was twelve males and females from each line in SD (N = 48 mice). Mice were selected at random, except that none of the mice chosen for the IGF-1 assay were full siblings. Only mice between 7 and 16 weeks old were used in the study.

Mice were kept in short day rooms (8 hours light : 16 hours dark, lights on at 0800 hours Eastern standard time) at the College of William and Mary Population Laboratory. Food and water were provided ad libitum. Mice were
caged separately and food in the hopper of each cage was weighed at one week
intervals for three weeks. To reduce inaccurate measures of food intake, mice
that chewed food that fell to the cage floor were excluded from analyses of food
intake. Body mass was also measured on the same days that food mass was
weighed. At the end of the three week period mice were culled (see below).

**Testosterone:** Sample size was 12 males of each line in each photoperiod
(N = 48 mice). Only mice between 7 and 16 weeks old were used in the study.
Mice were kept in short day (8 hours light : 16 hours dark, lights on at
0800 hours Eastern standard time) and long day rooms (16 hours light : 8 hours
dark, lights on at 0400 hours Eastern standard time) at the College of William
and Mary Population Laboratory. Food and water were provided *ad libitum*.
Mice were caged separately and body mass was measured at one week intervals
for three weeks.

**Estrogen:** Sample size was 12 females of each line in each photoperiod,
(N = 48 mice). Only mice between 7 and 16 weeks old were used in the study.
Living conditions of the mice were the same as in the testosterone experiment
explained above.

**Serum Collection:** Mice were fasted for 6 hours before blood collection in
order to reduce variation in feeding related hormones due to recent feeding. At
the time of blood collection, each mouse was taken quietly to a separate room
and individually anesthetized with isofluorane. While under deep anesthesia,
each mouse was euthanized and trunk blood was collected. For most blood
collections, blood was immediately tested for levels of glucose using a OneTouch
Ultra glucometer, which can measure levels of glucose from 20 to 600 mg/dL.
The accuracy of the glucometer was tested using a control solution. All
collections occurred between 1500 and 1700 hours.

Blood was allowed to clot about 30 minutes at 4° Celsius, subsequently
centrifuged, and serum was collected. All centrifuging was done at about 2,000 x
g on Beckman’s Microfuge B, Eppendorf’s miniSpin, and Fisher Scientific’s Micro
7. Serum that showed signs of contamination with dislodged or suspended blood
cells was respun. Serum was then frozen at -70° C for later use.

**Reproductive Organ Weights:** Immediately after blood collection, testes,
semenal vesicles, ovaries, and uteri were removed and weighed. Internal fluid
was removed from the seminal vesicles.

**Hormone Assays:** IGF-1 was assayed with a R&D Systems Quantikine
Mouse IGF-1 kit (Cat. MG-100). The assay had an intra-assay Coefficient of
Variation of 5.6%, 3.3%, and 4.1% for concentrations at 82, 269, and 921 pg/ml,
respectively. The detectable range for the assay standard curve was 31.2 pg/ml
to 2000 pg/ml. Prior to assay, serum was diluted 500 fold, and the levels
expressed in the results have been corrected for the dilution factor. All reagents
were used within the parameters set by the assay manual. The reconstituted
standard was used on the same day. Samples were run in singlet.

A Thermo-Scientific plate shaker, Model 4625, was used for all incubation
periods. The final reading of the IGF-1 plate was done using BioTek’s ELx800
Microplate Reader set at 450 nm with a correction filter of 540 nm. The assay’s
standard curve was built using a 4-P fit calculated by the plate reader.
Estrogen levels were quantified using a mouse ELISA kit made by Calbiotech. The ELISA was run by the UVA Center for Research in Reproduction. The estrogen assay had a sensitivity of 7.0 pg/mL, a reportable range of 3-300 pg/ml, an intra-assay CV of 4.4%, and an inter-assay CV of 7.8%. Samples were run in singlet.

Testosterone was quantified using an RIA made by Siemens (Coat-A-Count). This assay was also run by the UVA Center for Research in Reproduction. The testosterone assay had a sensitivity of 10 ng/dL, a reportable range of 7.4-740.8 ng/dL, an intra-assay coefficient of variation of 3.9%, and an inter-assay CV of 7.8%. Samples were run in singlet.

Validation of Testosterone and Estrogen Assays:

There were three tests used to test the validity of the UVA assays. These were especially important because we noticed that neither testosterone nor estrogen had many correlations with other reproductive measures. For these validation tests, pooled serum from groups of mice was vortexed before being aliquoted for assay.

Dilution curve: We developed a dilution curve to check for serum effects (Patrono 1987). Both the testosterone and the estrogen assays used have 100-fold ranges. Hence, we built a 100-fold dilution curve with five dilutions as follows: 1x, 1:1, 1:9, 1:49, and 1:99. We collected two pools of serum, one from 10 males and the other from 10 females. Each concentration of serum was assayed in sextuplicate. If serum affects were negligible, then we expected to see parallelism of the assay standard curve with the serum dilution curve (Patrono 1987; Billitti et al. 1998). Line, photoperiod, and age of mice did not matter for this test.

Positive and negative controls: We used positive and negative controls to verify that testosterone and estradiol were being measured (Patrono 1987). If other steroid hormones were cross-reacting with the antibodies of the assays, then we expected to find oddly high estradiol/testosterone in the negative controls. See “Analysis of Maturity” below for a description of how each group was selected. For both of the following experiments each group of serum samples was pooled and assayed in sextuplicate or quintuplicate.

Positive controls were 8 mature male and 8 mature female mice for the testosterone and estrogen assays, respectively. The mice were all raised in LD and most were Non-Responsive.

Negative controls were 8 immature male and 8 immature female mice for the estrogen and testosterone assays, respectively. The average age of immature mice was 43 days, and the oldest mouse was 87 days old. Most immature mice were Responsive mice that were raised in SD. These negative controls were expected to have low, but non-zero, levels of hormone.

Hormone spiking: In order to test whether a known amount of hormone can be added and recovered from serum (Patrono 1987), we collected and created two pooled serum samples from 24 mice. Each serum pool was split and half was placed in a tube spiked with hormone, while the other half was placed in a tube without hormone. To spike each tube we diluted a given amount of
hormone in ethanol. Ethanol was used because it dissolves steroids well and upon desiccation leaves crystals of steroids behind. Each dilution was made so that upon desiccation by SpeedVac we would have the desired amount of hormone, listed below. To allow hormone to dissolve fully into solution, each spiked serum pool was incubated at 4°C for at least 12 hours with its respective hormone. Line, photoperiod, and age did not matter for this test. Assays were run in sextuplicate or quintuplicate.

We collected 11 males and spiked half of their pooled serum (445 uL) with 13.06 ng of testosterone. This should have increased testosterone concentrations by about 293.48 ng/dL. The other half of the pooled serum was not spiked and was used as a control.

We collected 11 females and spiked half of their pooled serum (832.5 uL) with 8.04 ng of estradiol. This should have increased estrogen concentrations to 96.54 ng/dL. The other half of the pooled serum was not spiked and was used as a control.

**Statistics:**

Statistics were calculated using JMP on a Macintosh computer. A type 1 error of 5% (α=0.05) was used as a cut off value for all statistics, except determination of outliers. A type 1 error of 1% was used for determination of one outlier in our validation of the testosterone assay. For 1-way analyses, if we found that the two distributions had unequal variances, then we used a Welch ANOVA to correct for this. The following tests were used to determine if distributions had unequal variances: O'Brien, Brown-Forsythe, Levene, and Bartlett. The Bartlett test was excluded if the distributions were not normal.

For comparison of more than two means, 2-way ANOVA followed by a Tukey-Kramer test was used. Data on levels of testosterone were log transformed because of a highly skewed distribution. In addition to 2-way ANOVA, data on levels of testosterone were also analyzed by way of a non-parametric test, a 2-way Kruskal-Wallis. The Kruskal-Wallis test ranks data points from smallest to largest and then analyzes the data set that results like an ANOVA. Non-parametric tests, such as the Kruskal-Wallis test, do not require normality of data.

Matched pairs t-Tests were used to compare mature and immature mice of the assay validation experiment. Groups were given numeric identifiers in order to run the matched pairs analysis.

**Analysis of Maturity:**

When constructing positive and negative controls for the UVA assays, we assessed maturity by way of the following: testis index (length*width of a testis) and length of a seminal vesicle or greatest length of an ovary and width of the uterus. Using previous data sets (unpublished) of known maturities we were able to make rough estimates of maturity this way.
Results:
IGF-1:
The IGF-1 Assay had a standard curve with $R^2 = 1.000$. All four wells of controls were valid.

Males:
Testes mass of both R and NR males was correlated logarithmically with IGF-1 ($R = 0.75$ and 0.58, respectively; both $P<0.05$, figure 3). Seminal vesicles mass was correlated logarithmically to IGF-1 ($R = 0.41$, $P<0.05$ for the two lines collectively, figure 4).

Using a Welch ANOVA we found that no significant difference existed in mean levels of IGF-1 between NR and R males in SD ($P>0.46$, figure 5).

Figure 3: A) The correlation of testes mass and levels of IGF-1 of responsive males ($R = 0.75$, $P<0.05$). B) The correlation of testes mass and levels of IGF-1 of non-responsive males ($R = 0.58$, $P<0.05$). Both regressions use a log transform of testes mass, which tends to normalize the distributions.

Figure 4: Correlation of seminal vesicles mass and levels of IGF-1 hormone. ($R = 0.41$, $P< 0.05$). This analysis combines both the NR and R lines, and a log transform of seminal vesicle mass was used.
Figure 5: Results of ANOVA on levels of IGF-1 in males. The middle, horizontal line of the diamonds represents the mean. The tops and bottoms of the diamonds represent 95% confidence intervals.

Females:
Two NR females had levels of IGF-1 hormone that were above the standard curve and were assigned the highest values detectable by the curve.

The mass of ovaries was correlated linearly with levels of IGF-1 in R females ($R = 0.64$, $P < 0.05$, figure 6). The mass of uteri was correlated linearly with levels of IGF-1 in NR females ($R = 0.81$, $P < 0.01$, figure 7). A Welch ANOVA comparing the average levels of IGF-1 in NR and R females indicated no significant statistical difference ($P = 0.08$, figure 8).

Miscellaneous Measures:
Body weight was correlated with levels of IGF-1 in R females ($R = 0.73$, $P < 0.01$), and weight did not correlate significantly with levels of IGF-1 in the other groups. Food intake was not correlated with levels of IGF-1 in the males. The food intake of the females did correlate with IGF-1 levels ($R = 0.62$ and $0.76$ for NR and R respectively, both $P < 0.04$). Glucose did not differ significantly between the two lines ($P > 0.84$, figure 9), and only in NR females was glucose concentration correlated with levels of IGF-1 ($R = 0.76$, $P < 0.01$). In general, IGF-1 was correlated with the above measures related to metabolism only in females.

Figure 6: Correlation of ovary mass and levels of IGF-1 in R females. A linear regression has a $R = 0.64$ and $P < 0.05$. 
Figure 7: Correlation of uterine mass and levels of IGF-1 in NR females. A linear regression has a $R^2 = 0.81$ and $P<0.01$.

Figure 8: ANOVA comparing levels of IGF-1 in females of both lines. ($P = 0.08$).

Figure 9: Levels of glucose by line in SD (P>0.84). This test used the data of over 130 mice from various blood collections, including the IGF-1 collections.

**Testosterone:**
All of the testosterone radioimmunoassays had standard curves with $R \geq 0.9997$. All of the assays reported levels of testosterone of the quality controls within acceptable ranges.
In validating the testosterone RIA we found that mature males had significantly higher levels of testosterone than immature males (Matched Pairs: P<0.0001) and added testosterone was detectable. However, the testosterone spike was over-measured by the assay. The assay should have read values of about 300 ng/dL testosterone in the spiked serum, but instead all but one of the values were over the range of the assay (>800 ng/dL). The one value not over the range of the assay was below the range of the assay. This value was discarded after distribution analysis of all the spiked serum indicated that the one value was outside of a 99% confidence interval. Both the assay standard curve and our dilution curve have positive slopes (see figure 10), but the two are not parallel. In summary, the assay measured relative concentrations, but not absolute concentrations, of testosterone.

Using a generalized linear model (GLM) of log transformed levels of testosterone we tested the following effects: mouse line, testes mass, seminal vesicles mass, and photoperiod. Only photoperiod had a significant effect on levels of testosterone (P<0.02). We tested another GLM, except with the log transformed mass of testes as a response, and found that both line and photoperiod had significant effects on testes mass (both P<0.01). In summary, while both line and photoperiod were related to testes mass of the mice, only photoperiod was related to testosterone levels of the mice.

To determine if any differences existed in levels of testosterone between photoperiods, we log transformed the non-normal data on levels of testosterone. We then used both a 2-way ANOVA and a Tukey test to analyze the transformed data. A 2-way Kruskal-Wallis test was also used because the log-transformed data had unequal variances, which could have biased the Tukey test. All of the tests used resulted in the conclusions shown in figure 11. The means of both SD groups were significantly lower than the means of both LD groups (all P<0.05). There were no significant differences within photoperiods.

Figure 10: Linear Regressions of testosterone concentrations and dilutions in A) *Peromyscus* serum (P<0.01) and B) the assay standard curve (P = 0).

Note that A) has a different y-axis scale than B). A) has been presented thusly in order to show that the serum curve has a positive slope. The serum curve has a slope that is about one thousandth that of the standard curve. The y-axis range of the serum curve is 5—12 ng/dL compared to 5—1600 ng/dL for the standard curve.
Figure 11: Tukey analysis of male testosterone levels. Testosterone was log transformed to produce normality. LD = Long Day, SD = Short Day, NR = Non-Responsive, R = Responsive. The middle, horizontal line of the diamonds represents the mean. The tops and bottoms of the diamonds represent 95% confidence intervals. There are statistically significant differences between groups A and B, but not within each group.

**Estrogen:**
All quality controls of the assay were within acceptable values.

In validating the estrogen assay we found that mature females had significantly higher estrogen than immature females (P<0.0001), and the estrogen spike did increase the concentration of estrogen measured. However, the dilution curve did not parallel the assay standard curve (figure 12), and the dilution curve was of opposite slope compared to the assay standard curve. The dilution curve at first increased in estrogen concentration with an increase in dilution, then briefly decreased, and then rose again in measured estrogen concentration. We believe that a serum effect may have caused anomalies in the dilution curve. The estrogen spike caused an increase in estrogen, but we could not measure a specific increase in estrogen because of the range of the assay.

Using a GLM on levels of estrogen we found that none of the following had a significant relationship with estrogen: mouse line, photoperiod, ovary mass, and body weight. The GLM of estrogen did indicate that uterine mass was a significant effect (P<0.01, correlation: R=0.40). A GLM of uterine mass revealed that uterine mass was significantly related to photoperiod, line, and estrogen (all P<0.01). Therefore, while estrogen only related to uterine mass, uterine mass also related to line and photoperiod.

To determine if a difference existed in estrogen levels between groups we used 2-way ANOVA followed by a Tukey-Kramer test (all P-values>0.58, see figure 13). To verify that the Tukey analysis was not being biased by unequal variances, we tested whether or not there were any significant differences
between the variances of the groups (O’Brien, Brown-Forsythe, and Levene tests all P>0.38).

Figure 12: Estradiol Validation
These are dilution curves for *Peromyscus* serum and the assay standard, presented respectively (both P<0.05, R>0.90).

A) The *Peromyscus* serum dilution curve. B) The assay standard curve. Note that the serum curve has the opposite slope of the standard curve.
Figure 13: LD = Long Day, SD = Short Day, NR = Non-Responsive, R = Responsive. There is no statistically significant difference between any of the groups (all P-values > 0.58).

Table 1: Rough comparison of our results with results in the literature

<table>
<thead>
<tr>
<th>Species</th>
<th>Average IGF-1 (ng/mL)*</th>
<th>Average T (ng/dL)**</th>
<th>E (pg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Williamsburg Pero. leucopus</em></td>
<td>314</td>
<td>38</td>
<td>6.8 (averaged)</td>
</tr>
<tr>
<td>Pero. Leucopus</td>
<td>N/A</td>
<td>55</td>
<td>N/A</td>
</tr>
<tr>
<td>Pero. californicus</td>
<td>N/A</td>
<td>50</td>
<td>N/A</td>
</tr>
<tr>
<td>Rattus</td>
<td>137</td>
<td>215</td>
<td>145 to 2178</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>397</td>
<td>248</td>
<td>~5 to 23</td>
</tr>
</tbody>
</table>

*If both male and female levels were known then they have been averaged.
**If levels were known for both long and short days they have been averaged together.
1 Using measures of mice 8 weeks old from both SD and LD (Young et al. 1999).
2 (Oyegbile and Marler 2005)
3 (Yakar et al. 2002)
4 (Lee et al. 1975)
5 (Shaikh 1971)
6 (Miller et al. 2005)
7 (Zielinski and Vandenbergh 1993)
8 (Campbell et al. 1976; Nelson et al. 1981)
Conclusions and Discussion:
IGF-1:
This study supports previous studies that suggest IGF-1 is related to reproduction (Daftary and Gore 2005; Todd et al. 2007; Villalpando et al. 2008). The masses of the following reproductive organs correlated with levels of IGF-1: seminal vesicles, testes, uteri, and ovaries ($R^2$ values range between 0.1 and 0.65). Overall, these data suggest that levels of IGF-1 are related to life-history stages.

However, we found no evidence supporting our prediction that a difference would occur in IGF-1 concentrations between the mouse lines. We found no significant statistical differences in mean IGF-1 levels between the mouse lines, and we reject the evolutionary constraint hypothesis. This is not to say that the hypothesis is categorically wrong, but in this instance it seems not to apply. This does not mean that the evolutionary potential hypothesis is necessarily true. For the evolutionary potential hypothesis to be true we would have to find a difference in receptor expression or some other part of the IGF-1 hormone system.

This lack of a difference between the lines suggests that peripheral IGF-1 concentrations are unlikely to be the source of the phenotypic variation observed. This may have to do with the pleiotropic effects of IGF-1 on numerous other systems (Le Roith 1997). Because IGF-1 regulates the somatic growth of numerous tissues, selection on heritable variation in levels of IGF-1 to favor or disfavor reproductive development in SD could have maladaptive effects on somatic growth (see (Le Roith 1997) for possible effects). Therefore, the widespread effect of IGF-1 on multiple systems would likely constrain the evolution of changes in levels of IGF-1 (McGlothlin and Ketterson 2008).

In order to answer whether or not IGF-1 has some role in the observed phenotypic differences of our mouse lines, future experiments could investigate IGF-1 receptor expression and localization. Variation in IGF-1 receptor expression could enable NR testes, or other organs, to be stimulated by IGF-1 more strongly than R testes. If NR testes are more strongly stimulated by IGF-1 than R testes are, then this might explain the phenotypic difference between the lines.

Testosterone and Estrogen:
Our data lead us to reject the evolutionary constraint hypothesis for variation in levels of testosterone and estrogen within this population. No statistically significant difference in levels of testosterone occurs between the NR and R mouse lines within either LD or SD. The results of the estrogen assay are similar, but the results of estrogen assay need to be verified by another assay. This lack of a difference in hormone levels is preliminary evidence for the evolutionary potential hypothesis. However, more experiments must be considered.

Future experiments should discern whether central and/or peripheral receptor expression is the cause of the observed phenotype differences in the
NR and R mouse lines. Alternatively, the NR and R line may vary in their neuroanatomy controlling reproduction (Avigdor et al. 2005).

While there are no differences in hormone levels within photoperiods, a difference in levels of testosterone was found between photoperiods in the male mice. LD males have significantly more testosterone on average than either SD group. This suggests that there is a shift in the regulation of the testosterone hormone system from SD to LD in the males, for which reason we should continue using both SD and LD groups when analyzing testosterone.

**Caveats:**

Neither the testosterone assay nor the estrogen assay were validated fully. The testosterone assay did not measure a hormone spike accurately, which suggests that the testosterone assay measures relative concentrations of hormone. The estrogen assay appears to be biased by serum effects and has a small range. However, both assays do appear to be measuring their respective hormone.

Before concluding, there is another concern that should be addressed with respect to our selection lines. Phenotypic plasticity can be a concern in a selection line (Henderson 1997; Garland jr and Kelly 2006). We do not believe phenotypic plasticity is of great concern here for two reasons. First, the mice were kept in a constant environment. Second, the R line exhibits smaller testes and uterine measures in both SD and LD relative to NR peers (Heideman et al. 2005). Therefore, it is likely that the observed phenotypic differences between lines are genotypic in origin, rather than an effect of genotypic plasticity.

**Conclusions:**

This study sought to discover how three hormones affect life-history traits. We found that these hormones were variable in relation to life-history stages, such as reproductive measures or photoperiod response. However, the hormones did not differ between selection lines that vary heritably in the life history trait of photoperiod response.

Overall, this study suggests that peripheral hormone levels may be evolutionarily restricted by their multiple actions. If so, then the evolutionary constraint hypothesis may be rejected for this case. Rather than variation in levels of hormone affecting microevolutionary change, perhaps receptor systems more easily change and evolve. In this way receptor expression could change, causing a hormone to regulate an organ differently or to regulate an additional organ, without causing adverse pleiotropic effects. Future experiments should determine if a change in receptor expression could cause a difference in photoperiod response. If NR mice tend to have more receptors than R mice, then that would be evidence of the evolutionary potential hypothesis, depending on the receptor type.
Future Directions:

*IGF-1:*

The NR mice may express more IGF-1 receptors in their reproductive organs than R mice express. This would promote growth of NR reproductive organs, even with the same amount of IGF-1 hormone as R mice. Future studies should examine IGF-1 receptor expression in the following areas: testes, seminal vesicles, ovaries, uteri, and hypothalami. Hypothalamic studies should investigate co-localization of GnRH neurons with IGF-1 receptor.

It is also possible that the current study tested at a time of day in which levels of IGF-1 were not different between the lines. Consider a situation in which each mouse is cycling through different pulses of IGF-1. The NR line may have higher pulses, but this could have been obscured by the time at which serum was collected. To clarify this a study could measure IGF-1 hormone changes across time in individual mice. However, repeated measures are difficult in *Peromyscus leucopus* and a different route may be appropriate.

A growth hormone challenge to the two mouse lines should produce an increase in IGF-1 hormone that is the same between the lines. If higher and higher levels of growth hormone cause no change in levels of IGF-1 within the R line, but a continuing increase in levels of IGF-1 within the NR line, then it would suggest that the NR mice are capable of larger IGF-1 pulses than the R mice. Larger pulses would possibly explain why there is no overall difference in levels of IGF-1.

*Testosterone:*

The amount of testosterone in the NR line in short days must be sufficient for sperm production. However, since the R line in short days has about the same amount of testosterone, variation in the level of testosterone cannot explain the difference in reproductive phenotype of the two mouse lines. If the NR line has greater testosterone receptor expression in Sertoli cells, this might explain why the NR line has larger testes than the R line.

Alternatively, if the NR line has more Follicle Stimulating Hormone (FSH), then the excess FSH could promote testicular growth. It may also be worthwhile to measure male Luteinizing Hormone (LH) levels. A previous study found a difference in levels of LH between the females of the two lines (Heideman et al. 2010), and the current study made the assumption that male LH levels would be similarly different between lines. However, we currently do not know levels of LH in males, and should reexamine our assumptions.

*Estrogen:*

Another assay should be used to verify the current findings. The assay used in this study does not have the range required, and there are questions remaining regarding the validity of the assay. If another assay indicates that there is no difference between the lines or between photoperiods, then there are several other experiments that should be investigated.

Given that a previous study found higher levels of LH in NR females, we expected that the NR females would have higher levels of estrogen in the current
study (Heideman et al. 2010). The current lack of a difference between the lines could be a result of NR and R females maintaining a baseline level of estrogen following puberty. If so, then we would nonetheless expect to see higher estrogen in a female that is about to ovulate than a non-ovulating female. In order to determine whether or not we missed a difference in estrogen levels, a future study could categorize females within each line based on the number and type of follicles observed. If NR females with numerous follicles have more estrogen than all other females, then it would suggest that the current study missed the difference between the lines by averaging levels of estrogen.

Another possibility is that the brains of mice in the NR line may respond to increasing levels of estrogen more readily. This could occur if brains of NR mice express higher levels of estrogen receptor. Depending on how the estrogen receptor is expressed, increased expression of estrogen receptor could stimulate neurons that stimulate the ovulatory surge of LH and FSH. If NR females have a LH surge more readily than R females, then similar levels of estrogen could cause NR females to have higher levels of LH than R females. In order to test this, a study could measure levels of LH after adding varying amounts of estrogen to ovariectomized mice in SD. The same experiment should be done in LD, except R mice in LD should have a LH dosage response curve similar to NR mice in SD (see (Heideman et al. 2010) for data on levels of LH between each line of mice).
• Acknowledgements:
Paul Heideman for crucial guidance and support throughout the 3 years I have been working on these projects. Professor Paul Kaseloo for providing the plate reader, pipettes, the assay, and technical help. Professor Eric Bradley for being on my committee and for technical assistance. Professors John Griffin and Randolph Coleman for being on my committee. Professor Zwollo for help with preliminary assay data. Professors Mathias Leu and Helen Murphy for statistical assistance. Lydia Wright-Jackson for the care of the mice. Cori DeSanto and Madelyn Galimore for help with blood collections. David Berrigan, Salehin Rais, and Kathy Sharp for initial studies on the mice, which led to the studies of this thesis. Andrew Panakos and Leanne Pina for preliminary glucose and blood collection data.

This thesis received funding from the Howard Hughes Medical Institute Grant to the College of William and Mary. Jordan White was supported by a Charles Center Domestic Scholarship during one summer of research.

Literature Cited:


Appendix

Insulin:

Levels of insulin correlated with body mass in both R and NR mice (R = 0.70 and 0.89, respectively, both P<0.01, see figure A1 below). Insulin did not correlate with uterus masses, testes masses, seminal vesicle masses, ovary masses, or glucose levels at time of death. Between the lines there was no significant difference in levels of insulin (P>0.10, figure A2).

Figure A1: Linear regressions containing both females and males (both P<0.01).

Figure A2: The error bars represent 95% CI's (P>0.10). Each bar contains both males and females of the given line.

Corticosterone:

Levels of corticosterone correlated negatively with testes mass in NR mice in SD (R = 0.99 figure A3). The data shown exclude most of the mice in the experiment because they had collection times over 3 minutes or corticosterone values well beyond the standard curve of the assay. All of this data needs to be reworked because most mice had corticosterone values over the detectable range of the hormone assay. Future experiments should collect one or two mice at each collection. The mice should be anesthetized immediately outside of their cage rooms, rather than being carried all the way to the surgery room. Blood should be collected immediately outside of the cage room. This should speed up blood collection and minimize the possibility of a stressed mouse being collected.
Note that corticosterone levels tend to change noticeably ~3 minutes after a stressful stimulus.

Figure A3:
Logarithmic regression was used on NR male data from SD (R = 0.99, P<0.01). The R male data shows a non-significant correlation.

Ghrelin:
Levels of ghrelin were not correlated with any other measure. The assay requires validation, but other hormone systems are more likely sources of the differences observed in the mice. Furthermore, the entire reason we investigated levels of ghrelin was because the NR line ate more food than the R line. This does not seem to be the case in these lines at present, after 6 generation without selection. Therefore, this hormone seems to have little promise for further investigation.

Leptin:
Levels of leptin correlated with body mass collectively in all mice grouped together (R = 0.88, P<0.01 figure A4). Leptin also correlated to varying degrees with a few reproductive measures, including testes and seminal vesicles (all 0.71>R>0.32). However, the assay used was not sensitive enough to detect the lowest leptin values, which mostly were from R mice. It is possible that a difference exists between the lines in level of leptin, but this assay needs to be run again. Future experiments should try to find a more sensitive assay, possibly a RIA offered by an assay facility. Additionally, larger and more diverse samples should be used. Unintentionally, I randomly selected R mice that were mostly very skinny. Because fat cells are required to produce leptin, this may have caused a skewing in the data set.
Figure A4: This regression compiles all of the mice ($R = 0.88$, $P<0.01$).

Figure A5: Mean levels of leptin were not significantly different ($P>0.05$). The error bars are 95% CI's.