



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

ENES JURIQUILLA

DINÁMICAS DE EXPRESIÓN Y EVOLUTIVAS DE
UN CIRCUITO DE REGULACIÓN GÉNICA CON
RETROALIMENTACIÓN POSITIVA Y NEGATIVA

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

Licenciada en Ciencias Genómicas

PRESENTA:

Frida Sophia Orozco Ruiz

TUTOR:

Dra. Mariana Gómez Schiavon

Juriquilla, Querétaro, 2022



LIIG-UNAM



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis papás.

Agradecimientos

Quiero agradecerle en especial a mi tutora y directora de tesis, la Dra. Mariana Gómez Schiavon. Agradezco todas las oportunidades y apoyo que me ha brindado, por siempre estar dispuesta escuchar y discutir ideas, y por impulsarme y prepararme para seguir con el resto de mi camino académico. Sobre todo le agradezco la confianza que ha tenido en mi y por hacer de mi cuarto año de licenciatura una etapa muy feliz. Con mucho cariño le reconozco su excelente trabajo como tutora.

También quiero agradecer al resto de mi jurado conformado por el Dr. Pedro Eduardo Miramontes Vidal, el Dr. Vicente Diego Ortega Del Vecchyo, la Dra. Ayari Fuentes Hernández y el Dr. Rafael Peña Miller por sus comentarios, apoyo y disponibilidad durante mi proceso de titulación. En particular a Pedro y a Diego, que ambos fueron mis maestros durante la licenciatura, les agradezco por contagiarme su entusiasmo y compartirme sus conocimientos sobre los sistemas complejos y sobre evolución, respectivamente.

Agradezco el apoyo y disponibilidad de Jair García Sotelo del Laboratorio Internacional de Investigación sobre el Genoma Humano (LIIGH UNAM) y de Luis Alberto Aguilar del Laboratorio Nacional de Visualización Científica Avanzada (LAVIS UNAM). El apoyo técnico que me brindaron fue indispensable para el desarrollo de este proyecto.

Agradezco al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT UNAM) (Clave del proyecto: IA204022) por la beca que recibí durante mi último semestre de estudios y en mi proceso de titulación.

Agradezco a la Dra. Margareta Boege por sus muy útiles comentarios sobre mi tesis y por siempre estar al pendiente de todos en la licenciatura. Finalmente, quiero agradecer a mi familia y amigos. A mis papás, por a través de su esfuerzo darme la oportunidad de estudiar una carrera universitaria y por el apoyo que me han dado para realizar cualquier objetivo que se me ocurriera. A mis amigos por el cariño y apoyo. En particular Moni y Erick, mis compañeros y amigos del salón, que siempre estaban dispuestos a discutir, ayudar y escuchar (en relación a la ciencia y, más importante, en todo lo demás).

Resumen

Los circuitos de regulación génica son capaces de generar dinámicas de expresión que emergen de sus interacciones; dos de estas dinámicas son las oscilaciones y el *switching* epigenético (i.e. sistemas biestables que en presencia de ruido bioquímico pueden cambiar de un nivel de expresión estable a otro). En particular, estos dos comportamientos requieren de circuitos con retroalimentación negativa y positiva, respectivamente. Se ha propuesto que el surgimiento y permanencia de estos circuitos puede darse bajo ciertas condiciones evolutivas que favorezcan a las dinámicas que son capaces de generar. El objetivo de este trabajo es explorar las dinámicas de expresión de un circuito con retroalimentación positiva y negativa, en especial las oscilaciones y la biestabilidad, el efecto del ruido bioquímico en ellas, y el potencial de las oscilaciones y el *switching* epigenético como estrategias para lidiar con ambientes fluctuantes cuando estas están en competencia entre ellas y con la adaptación genética. Esto se investigó a través del desarrollo de un modelo de ecuaciones diferenciales, simulaciones estocásticas del modelo y de experimentos evolutivos in silico. Se encontró que el modelo desarrollado puede presentar los comportamientos de interés además de otras dinámicas que solo son posibles en presencia del ruido bioquímico: oscilaciones inducidas por ruido y sistemas excitables, que mientras que en simulaciones deterministas llegan a un estado estacionario en las simulaciones estocásticas presentan un comportamiento más dinámico. En las simulaciones evolutivas se identificó que las oscilaciones, el *switching epigenético* y las oscilaciones inducidas por ruido pueden surgir como estrategias para lidiar con ambientes fluctuantes.

Índice general

Índice de figuras	ix
Índice de tablas	xv
1. Introducción	1
2. Marco Teórico	3
2.1. Circuitos de regulación génica	3
2.1.1. Regulación génica	3
2.1.2. Motivos y propiedades emergentes	4
2.1.3. Modelos matemáticos de circuitos de regulación génica	6
2.1.4. Biestabilidad	8
2.1.5. Oscilaciones	10
2.2. Ruido Bioquímico	11
2.2.1. Modelos estocásticos	12
2.2.2. <i>State-switching</i>	13
2.2.3. Oscilaciones ante el ruido bioquímico	14
2.3. Adaptación ante ambientes fluctuantes	16
2.3.1. Variación genética	16
2.3.2. Variación no genética	17
2.3.3. Condiciones ambientales que favorecen a las estrategias adaptativas	17
3. Metodología	19
3.1. Desarrollo del modelo	19
3.1.1. Simulaciones Deterministas del sistema	23
3.1.2. Análisis de estabilidad	23
3.1.3. Exploración del espacio de parámetros	24
3.1.3.1. Exploración basada en fenotipos	24
3.1.3.2. Exploración basada en parámetros	26
3.2. Simulaciones Estocásticas del sistema	27
3.3. Modelo evolutivo	27
3.3.1. Antecedentes	28
3.3.2. Descripción	28

3.3.3. Diseño Experimental de las Simulaciones Evolutivas	33
3.3.4. Análisis de las simulaciones evolutivas	35
4. Dinámicas del circuito de regulación génica con retroalimentación positiva y negativa	37
4.1. Solución oscilatoria	37
4.2. Solución biestable	38
4.3. Caracterización del espacio de parámetros	38
4.4. Discusión	42
5. Efecto del Ruido Bioquímico en las dinámicas del circuito	43
5.1. Las oscilaciones son robustas ante el ruido bioquímico	43
5.2. Oscilaciones inducidas por ruido bioquímico	43
5.3. <i>Switching</i> biestable	45
5.4. Sistema excitable inducido por ruido	47
5.5. Discusión	47
6. Dinámicas evolutivas del circuito ante un ambiente fluctuante	49
6.1. Las poblaciones logran adaptarse ante ambientes fluctuantes	49
6.2. Estrategias para lidiar con ambientes fluctuantes	52
6.2.1. Dinámicas emergentes del circuito como estrategias adaptativas .	54
6.2.1.1. Oscilaciones	54
6.2.1.2. Biestabilidad - <i>State Switching</i>	55
6.2.1.3. Oscilaciones inducidas por ruido	56
6.2.2. Adaptación genética como estrategia adaptativa	58
6.3. Discusión	59
7. Conclusiones y Perspectivas	63
A. Códigos	65
A.1. Análisis de estabilidad de los estados estacionarios	65
A.2. EvoDynamics_OscillatoryModel.cpp (Selección sobre B)	71
A.3. GeneNetwork_OscillatoryModel.h (Selección sobre B)	78
Bibliografía	91

Índice de figuras

2.1. Ejemplos de motivos de redes de regulación génica - Diferentes formas de obtener retroalimentación positiva (A) y retroalimentación negativa (B). Un número impar de interacciones de regulación negativas hacen que el circuito presente retroalimentación negativa. Las flechas que terminan en punta indican activación de la transcripción mientras que las otras indican represión.	5
2.2. Función de Hill para regulación transcripcional positiva - Con parámetros $K = 80$, $K_D = 40$ y n con valores según se indica en la figura. El parámetro K_D es la concentración de X requerida para que la función de síntesis llegue a la mitad de su síntesis máxima.	7
2.3. Biestabilidad - Tasa de cambio de X (ecuación 2.5) con respecto a X con parámetros $K = 80$, $\alpha = 0.25$, $K_D = 50$, $\gamma = 1$ y n según se indica en la figura. Con suficiente no linealidad ($n = 10$) el sistema es biestable mientras que con $n = 1$ el sistema es monostable (i.e. un solo punto de equilibrio). Los puntos negros representan puntos de equilibrio atractores y el punto blanco el punto de equilibrio repulsor.	9
2.4. Oscilador de Goodwin - Simulación determinista del oscilador de Goodwin con parámetros: $K_1 = K_D = K_3 = K_5 = 1$, $K_2 = K_4 = 0.1$, $K_6 = 0.11$ y $n = 10$; con condiciones iniciales $X = 0$, $Y = 0.2$ y $Z = 2.5$. En la izquierda se muestran las concentraciones de las proteína y mRNA con respecto al tiempo. En la derecha el ciclo límite.	11
2.5. Switch Epigenético - Simulación estocástica del modelo descrito en la sección 2.1.3, ecuación (2.5). En la izquierda se muestra la serie de tiempo de la simulación del número de moléculas de la proteína X ; a la derecha su histograma. Las modas de la distribución corresponden a los estados estacionarios del sistema: $X = 20$ y $X = 80$	14
2.6. Circuitos centrales de osciladores robustos - Las tres estructuras más robustas ante perturbaciones en el espacio de parámetros. A) <i>Repressilator</i> . B) Retroalimentación negativa con retraso. C) Activador-represor.	15

3.1. Diagrama del circuito de retroalimentación positiva y negativa - El gen A se auto-activa al mismo tiempo que activa al gen B, quien reprime a A.	20
3.2. Nullclines - Parámetros de la <i>nullcline</i> de A: $K_1 = 100$, $\alpha = 0.5$, $K_A = 20$, $K_B = 20$, $n_A = 5$, $n_B = 5$, $K_2 = 1$. Parámetros de la <i>nullcline</i> de B: $K_3 = 70$, $\beta = 0.01$, $K = 25$, $n = 5$, $K_4 = 1$	24
3.3. Esquema de la dinámica ambiental con selección sobre la proteína A - El ambiente fluctúa cada $\frac{1}{f}$ generaciones entre $E = \{H, L\}$. El número de generaciones que permanecen en el mismo ambiente corresponden a una época. Dos épocas en ambientes distintos ocurren cada $\frac{2}{f}$ y corresponden a un ciclo ambiental. En cada ambiente se selecciona para un número de proteínas distinto, correspondientes a $A^{(H)} = 80$ y $A^{(L)} = 20$, cuando se está en el ambiente alto (H) y bajo (L), respectivamente.	29
3.4. Esquema del ciclo generacional con selección sobre la proteína A - Cada integrante de la población de tamaño N tiene el circuito de regulación génica y un genotipo propio. En cada individuo se simulan las dinámicas de expresión estocásticas del circuito. Se evalúa el valor de la proteína A al final de la simulación y se le asigna un valor de adaptación $w(A)$ a la célula, en el que el valor óptimo de A depende del ambiente. La siguiente generación se obtiene a través de selección de torneo (donde la presión selectiva está relacionada con el número de miembros del torneo S_t) de acuerdo al valor de $w(A)$. Las células de la siguiente generación tienen una probabilidad u de mutar en cuatro posibles escenarios y con un paso mutacional M	30
3.5. Genotipos iniciales - Parámetros para los genotipos iniciales en la selección sobre A (izquierda): (1) $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 50$, $n = 4$, $K_4 = 1$. (2) $K_1 = 100$, $\alpha = 0.25$, $K_A = 30$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 40$, $n = 4$, $K_4 = 1$ (3) $K_1 = 100$, $\alpha = 0.25$, $K_A = 6$, $K_B = 10$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 120$, $\beta = 0.01$, $K = 40$, $n = 4$, $K_4 = 1$. (4) $K_1 = 84$, $\alpha = 0.6$, $K_A = 25$, $K_B = 15$, $n_A = 16$, $n_B = 7$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 24$, $n = 2$, $K_4 = 1$. Parámetros para los genotipos iniciales en la selección sobre B (derecha): (1) $K_1 = 50$, $\alpha = 0.5$, $K_A = 3$, $K_B = 10$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 20$, $n = 5$, $K_4 = 1$. (2) $K_1 = 50$, $\alpha = 0.5$, $K_A = 30$, $K_B = 20$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 50$, $\beta = 0.01$, $K = 20$, $n = 5$, $K_4 = 1$ (3) $K_1 = 80$, $\alpha = 0.25$, $K_A = 10$, $K_B = 15$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.01$, $K = 20$, $n = 3$, $K_4 = 1$. (4) $K_1 = 84$, $\alpha = 0.6$, $K_A = 25$, $K_B = 15$, $n_A = 16$, $n_B = 7$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 24$, $n = 2$, $K_4 = 1$	34

3.6. Diagrama de la elección de individuos que se analizan para caracterizar las estrategias adaptativas - Las células marcadas con una estrella son aquellas cuyas dinámicas se caracteriza, pues son las de la última generación de cada época del primer ciclo de aquellos linajes que sobreviven dos ciclos ambientales.	36
4.1. Dinámica oscilatoria - Simulación determinista del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 30$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (izquierda). Plano fase con las <i>nullclines</i> del sistema (derecha).	38
4.2. Dinámica biestable - Simulación determinista del circuito con parámetros: $K_1 = 80$, $\alpha = 0.5$, $K_A = 40$, $K_B = 30$, $n_A = 8$, $n_B = 6$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.25$, $K = 20$, $n = 3$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (líneas punteadas) y $A = 60$ y $B = 30$ (líneas sólidas) (izquierda). Plano fase con las <i>nullclines</i> del sistema (derecha).	39
4.3. Gráfica de estabilidad de acuerdo a DST - Estabilidad de las regiones de parámetros de acuerdo a cambios en K y K_3 . Las clasificaciones de estabilidad están descritas en el texto principal. El conjunto de parámetros referencia para esta figura: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; este conjunto se encuentra en la región 2 de la figura.	40
4.4. Gráfica de estabilidad de acuerdo a la exploración basada en parámetros - Muestreo de 5,000 conjuntos de parámetros. El rango de valores de K y K_3 es $[0,3]$ en escala logarítmica. Parámetros fijos: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $\beta = 0.01$, $n = 4$, $K_4 = 1$	41
4.5. Dinámica en la región (0,1,2) - Simulación determinista del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (izquierda). Plano fase con las <i>nullclines</i> del sistema (derecha).	42
5.1. Dinámica oscilatoria con ruido bioquímico - Simulación estocástica del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 30$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$	44
5.2. Oscilaciones inducidas por ruido bioquímico - Simulaciones estocástica (izquierda) y determinista (derecha) del circuito con parámetros: $K_1 = 90$, $\alpha = 0.5$, $K_A = 30$, $K_B = 40$, $n_A = 8$, $n_B = 5$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.01$, $K = 25$, $n = 15$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$	44

5.3.	Oscilaciones cerca de un punto de bifurcación - Simulación determinista del circuito con parámetros: $K_1 = 90$, $\alpha = 0.5$, $K_A = 30$, $K_B = 40$, $n_A = 8$, $n_B = 5$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.01$, $K = 28$, $n = 15$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$	45
5.4.	Switch biestable - Simulación estocástica del circuito con parámetros: $K_1 = 80$, $\alpha = 0.5$, $K_A = 40$, $K_B = 30$, $n_A = 8$, $n_B = 6$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.25$, $K = 20$, $n = 3$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (figura superior). Histogramas del número de moléculas de ambas proteínas (figura inferior).	46
5.5.	Simulación estocástica del Fenotipo (0,1,2) - Simulación estocástica del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$	47
6.1.	Adaptación promedio de las poblaciones - Valores de adaptación máxima del promedio de las poblaciones en los últimos 10 ciclos para cada experimento (grupos horizontales) y sus respectivos genotipos iniciales (grupos verticales) bajo las diferentes condiciones simuladas variando los parámetros M (filas) y f (columnas). Los genotipos iniciales son los siguientes: (1) estado estacionario alto en $P \sim 80$, (2) estado estacionario bajo en $P \sim 20$, (3) oscilaciones con picos y valles en $P \sim 80$ y $P \sim 20$, respectivamente y (4) biestabilidad con estados estacionarios en $P \sim 80$ y $P \sim 20$. Para las simulaciones que seleccionan sobre A, P corresponde a A, mientras que para las simulaciones que seleccionan sobre B, P corresponde a B. Los valores de adaptación se obtuvieron al promediar el valor máximo de adaptación promedio de la población en el ambiente bajo con el valor máximo de adaptación promedio de la población en el ambiente bajo, ambos dentro de los últimos 10 ciclos ambientales.	50
6.2.	Casos en los que las poblaciones logran adaptarse - Promedio poblacional de adaptación (w) del primer y último ciclo ambiental; la población inicia en el ambiente bajo y después de $1/f$ generaciones cambia al ambiente alto. Histogramas de los promedios del número de proteínas en las 10,000 generaciones simuladas. Caso de selección sobre A con genotipo inicial 3, $f = 0.02$ y $M = 3$ (superior). Caso de selección sobre B con genotipo inicial 2, $f = 0.2$ y $M = 2$ (inferior).	51
6.3.	Estrategias de los linajes - Proporción de cada estrategia utilizada bajo las diferentes condiciones simuladas variando los parámetros M (filas) y f (columnas); agrupados según las estrategias descritas en el texto principal y los genotipos iniciales.	53

<p>6.4. Oscilaciones como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo de uno de los linajes que utiliza OSC como estrategia, con parámetros: $K_1 = 160.08$, $\alpha = 0.99$, $K_A = 5.26$, $K_B = 16.61$, $n_A = 7.02$, $n_B = 15.38$, $K_2 = 1$, $K_3 = 160.24$, $\beta = 0.006$, $K = 37.43$, $n = 1.77$, $K_4 = 1$; se muestran dos decimales (derecha).</p>	55
<p>6.5. Biestabilidad-<i>State Switching</i> como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo de uno de los linajes que utiliza BIS como estrategia, con parámetros: $K_1 = 70.8$, $\alpha = 0.36$, $K_A = 18.14$, $K_B = 17.58$, $n_A = 11.91$, $n_B = 6.04$, $K_2 = 1$, $K_3 = 82.46$, $\beta = 0.14$, $K = 17.4$, $n = 2.23$, $K_4 = 1$, se muestran dos decimales (derecha).</p>	56
<p>6.6. Oscilaciones inducidas por ruido como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 38.58$, $K_B = 8.69$, $n_A = 5.45$, $n_B = 1.07$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.007$, $K = 69.45$, $n = 16$, $K_4 = 1$, se muestran dos decimales (derecha).</p>	57
<p>6.7. Adaptación genética como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo del ambiente E=H con parámetros: $K_1 = 318.35$, $\alpha = 0.06$, $K_A = 0.01$, $K_B = 0.03$, $n_A = 0.03$, $n_B = 0.2$, $K_2 = 1$, $K_3 = 113.41$, $\beta = 0.035$, $K = 17.93$, $n = 0.12$, $K_4 = 1$. Simulación estocástica del genotipo del ambiente E=L con parámetros: $K_1 = 439.61$, $\alpha = 0.03$, $K_A = 0.01$, $K_B = 0.09$, $n_A = 0.02$, $n_B = 0.5$, $K_2 = 1$, $K_3 = 113.41$, $\beta = 0.035$, $K = 17.93$, $n = 0.12$, $K_4 = 1$, se muestran dos decimales (derecha).</p>	58

Índice de tablas

3.1. Reacciones del modelo	27
3.2. Escenarios mutacionales	32
3.3. Parámetros utilizados para las simulaciones evolutivas	35
4.1. Fenotipos del sistema	39

Introducción

La regulación génica requiere de complejas interacciones entre genes y proteínas que le permiten a las células controlar la expresión de sus genes, de manera que puedan realizar los procesos necesarios para sobrevivir y funcionar adecuadamente. Estas interacciones forman circuitos de regulación génica. Para cumplir sus diversas funciones, las células deben lograr que sus genes se expresen en diferentes patrones, los cuales están fuertemente determinados por la estructura de sus circuitos de regulación. Dos de estas dinámicas de expresión son las oscilaciones y el *state switching* o *switching* epigenético. Las oscilaciones son utilizadas para generar variaciones periódicas en los niveles de expresión de los genes y requieren de un circuito de regulación con retroalimentación negativa (28) (e.g. un gen que inhibe su propia expresión). Por otra parte, a través de procesamiento de señales o utilizando el ruido bioquímico, el *state switching* le permite a las células variar la expresión de los genes entre dos estados estables (i.e. biestabilidad); para ello el circuito requiere de retroalimentación positiva (36) (e.g. un gen que activa su propia expresión).

A los circuitos con funciones específicas asociadas y que además tienen un bajo número de componentes en su estructura se les conoce como motivos de redes de regulación (1). Por dichas características y por ser recurrentes dentro de las redes de regulación se sugiere que pudieran ser resultado de evolución convergente, pues se ha visto que en muchos casos los genes del circuito no están relacionados y no parecen ser resultado de la duplicación de circuitos ancestrales (1). Por ello, se cree que los motivos pudieron haber surgido y sido seleccionados varias veces porque las dinámicas que son capaces de generar resultan ventajosas bajo ciertas condiciones evolutivas. Entender las condiciones que favorecen dichas dinámicas continúa siendo una pregunta inconclusa y relevante en la biología evolutiva y de sistemas.

Recientemente y a través de simulaciones evolutivas, se ha observado que en ambientes que fluctúan entre dos estados a altas frecuencias el *state switching* o *switching* epigenético, que permite un cambio espontáneo entre estados, puede ser una estrategia ventajosa en comparación con la adaptación genética (16). En cuanto a las oscilacio-

1. INTRODUCCIÓN

nes, aunque se ha observado que ciertos circuitos capaces de generar la dinámica están bien conservados tanto en sistemas naturales como sintéticos, poco se sabe sobre las condiciones ambientales que favorecen su selección.

El objetivo de este trabajo es explorar las dinámicas evolutivas y de expresión, en particular la biestable y la oscilatoria, que puedan generarse de un circuito de regulación génica con retroalimentación positiva y negativa. Estas dinámicas se exploraron desde tres perspectivas. Primero, se desarrolló un modelo de ecuaciones diferenciales del circuito descrito y a través de simulaciones numéricas y de análisis de estabilidad se caracterizaron regiones del espacio de parámetros que presentaban oscilaciones y biestabilidad. Segundo, se realizaron simulaciones estocásticas del modelo para estudiar el efecto del ruido bioquímico en las dinámicas del circuito. Tercero, a través de simulaciones evolutivas que toman en cuenta la expresión estocástica de los genes del circuito, se exploró el potencial de los *switches* epigenéticos y de las oscilaciones como estrategias que las poblaciones pudieran utilizar para lidiar con ambientes fluctuantes. Probando diversas condiciones ambientales se buscó ganar un entendimiento de las condiciones que pudieran favorecer a cada estrategia mientras se encuentran en competencia entre ellas y con la adaptación genética.

Este trabajo está dividido en 7 capítulos. En el Capítulo 2 se encuentra el Marco Teórico; aquí se presentan las ideas de los circuitos de regulación génica, biestabilidad, oscilaciones y las bases de su modelado matemático; el ruido bioquímico, su rol funcional y modelos estocásticos, y por último una perspectiva general de las estrategias que los organismos pueden utilizar para lidiar con ambientes fluctuantes. En el Capítulo 3 se describe la metodología del trabajo. Los Capítulos 4, 5 y 6 corresponden a las secciones de resultados de cada una de las tres perspectivas descritas: caracterización de las dinámicas y del espacio de parámetros del modelo, el efecto del ruido bioquímico en las dinámicas y los resultados de las simulaciones evolutivas del circuito. Finalmente, las Conclusiones y Perspectivas se encuentran en el último capítulo.

2.1. Circuitos de regulación génica

2.1.1. Regulación génica

La regulación génica es determinante en la funcionalidad e identidad celular. El genoma de un organismo constituye miles de genes, pero para el funcionamiento correcto de las células no todos ellos pueden expresarse al mismo tiempo o en las mismas cantidades. Por ejemplo, las proteínas y transcritos (y sus concentraciones) de una célula sanguínea son diferentes a las de una neurona, aún cuando ambas poseen el mismo genoma. La regulación génica, a través de complejas interacciones entre genes y proteínas, permite a las células controlar la expresión de sus genes en términos de cantidad, tiempo y localización. La expresión de un gen puede ser regulada en distintas etapas (e.g. evitar la traducción de los RNA mensajeros), pero la forma más común es al inicio de la transcripción: mientras que solo hay una o dos copias por gen en una célula, miles de moléculas de RNA pueden ser sintetizadas a partir de un solo gen; evitar que el gen se transcriba puede resultar más sencillo y energéticamente más eficiente (39).

El proceso de transcripción funciona de la siguiente manera: en ausencia de proteínas reguladoras, la enzima RNA polimerasa puede unirse al promotor de un gen y comenzar a sintetizar moléculas de RNA, pero su afinidad por el promotor suele ser débil y por lo tanto solo se une ocasionalmente, haciendo que la expresión del gen llegue únicamente a un nivel bajo llamado basal (39). Las proteínas reguladoras, también llamadas factores de transcripción (TFs por sus siglas en inglés), reconocen regiones del DNA cerca de los genes para incrementar su expresión (TFs conocidos como activadores) o para inhibirla (TFs conocidos como represores).

Los activadores suelen tener afinidad a una región cercana al promotor. Al unírsele también interactúan con la RNA polimerasa y la atraen al promotor del gen para que se inicie la transcripción (39). De esta manera el activador facilita la unión de la en-

zima y el promotor a través del fenómeno de *cooperatividad* (i.e. que un sustrato se una a una enzima facilita el resto del proceso) (39). Por otra parte, los TFs represores simplemente deben unirse a la región de DNA del gen con la que interactúa la RNA polimerasa, evitando así la transcripción (39). También existen activadores y represores que realizan regulación alostérica, de manera que promueven un cambio conformacional en el DNA o en la RNA polimerasa para cumplir su función (39).

Además de los mecanismos anteriores, en células eucariontes la regulación suele ser más compleja por la presencia de histonas y *enhancers*. Las primeras dificultan la expresión génica en ausencia de factores de transcripción y modificaciones en ellas (e.g. metilación) están asociadas con cambios en la expresión (39). Las segundas son regiones de DNA donde factores de transcripción pueden unirse y promover la transcripción de genes (39).

2.1.2. Motivos y propiedades emergentes

Las complejas interacciones que surgen de los procesos descritos en la sección anterior forman circuitos y redes de regulación génica. Los circuitos y redes están formados por nodos (e.g. genes) y conexiones que representan el tipo de regulación entre ellos (i.e. positiva o negativa). Dentro de las redes de regulación existen circuitos recurrentes (i.e. patrones de interacciones) llamados motivos (1). Se ha teorizado que a lo largo de la evolución los motivos han surgido y han sido seleccionados múltiples veces al ser estructuras simples (por su bajo número de nodos) con funciones específicas de procesamiento de señales que podrían ser ventajosas en ciertos ambientes.

Dentro de los motivos más comunes se encuentran la retroalimentación positiva y negativa, esquematizadas en la Figura 2.1. La retroalimentación positiva se da cuando un TF (X en la Figura 2.1) activa su propia síntesis y se caracteriza por tiempos de respuesta lentos y porque suele aumentar la variación de los niveles de expresión entre las células. Lo primero se debe a que el gen comienza a expresarse en bajas concentraciones y con una velocidad de síntesis lenta hasta que llega a cierta concentración, conocida como el umbral de activación, a partir del cual la síntesis se vuelve rápida y la proteína llega a su estado estacionario; el tiempo de respuesta del motivo suele ser más lento que el de una regulación simple (i.e. cuando el TF regula a otro gen que no es el mismo). Cuando la regulación es débil, la variación suele ser mayor que en la regulación simple y cuando es fuerte puede llegar a generar distribuciones bimodales. Por otra parte, la retroalimentación negativa se da cuando un TF (X en la Figura 2.1) inhibe su propia síntesis y tiene las características opuestas de la retroalimentación positiva. Si inicialmente su expresión es alta (a través de un promotor fuerte) rápidamente llega al umbral de represión y a su estado estacionario, lo que le permite tener menor variación.

Los dos motivos descritos son la base para las dinámicas de interés de este trabajo: la retroalimentación positiva puede presentar biestabilidad y la negativa, si cuenta con

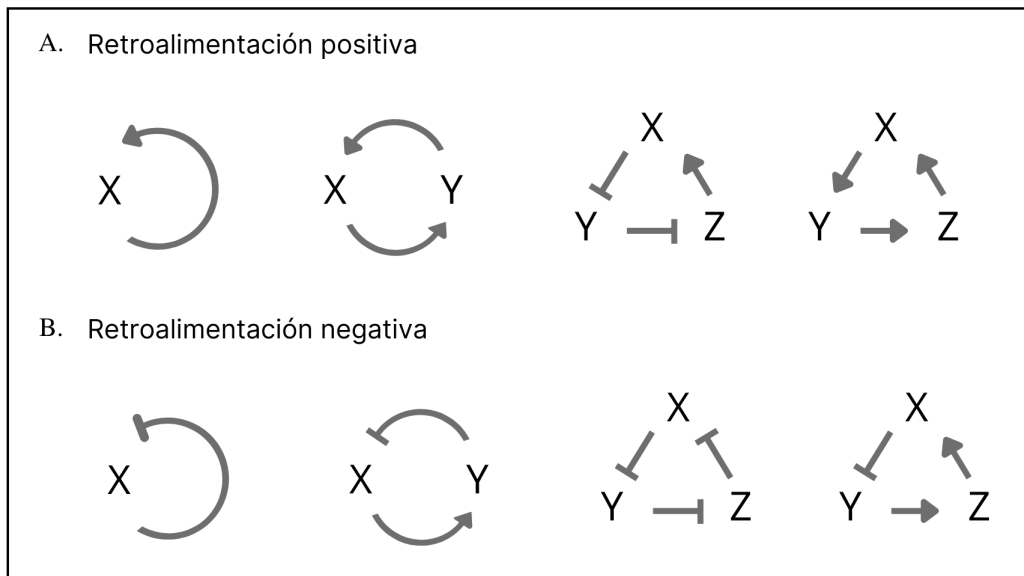


Figura 2.1: Ejemplos de motivos de redes de regulación génica - Diferentes formas de obtener retroalimentación positiva (**A**) y retroalimentación negativa (**B**). Un número impar de interacciones de regulación negativas hacen que el circuito presente retroalimentación negativa. Las flechas que terminan en punta indican activación de la transcripción mientras que las otras indican represión.

suficiente retraso temporal, oscilaciones. En general, a estos y a otros comportamientos que surgen en las redes se les conoce como propiedades emergentes de los circuitos de regulación génica (15), pues resultan de las interacciones que componen a los circuitos y no se podrían observar al estudiar individualmente los genes o proteínas que los conforman.

2.1.3. Modelos matemáticos de circuitos de regulación génica

Existen diferentes maneras de estudiar las interacciones que conforman a los circuitos de regulación génica. Una de ellas es a través de modelos matemáticos. Estos son abstracciones de los complejos sistemas biológicos y pueden servir como herramientas para obtener un mejor entendimiento de ellos, recapitulando sus propiedades. Modelos de pocas variables pueden servir para entender los principios que permiten el funcionamiento del sistema (e.g. por qué un motivo de retroalimentación negativa puede presentar oscilaciones); por otra parte, modelos más complejos suelen utilizarse para predecir comportamientos específicos (e.g. como la pérdida de un gen afecta el funcionamiento de una red de regulación de varios componentes). Aunque hay diferentes tipos de modelos, para entender las dinámicas de los circuitos de regulación génica, los sistemas de ecuaciones diferenciales han sido ampliamente utilizados. A continuación, se desarrolla un modelo de ecuaciones diferenciales del motivo de retroalimentación positiva (Figura 2.1 A) con el fin de plantear los principios detrás de este tipo de representación de la regulación génica.

La tasa de cambio de la concentración del factor de transcripción X , quién es un activador de su propia síntesis, está dada por la siguiente ecuación:

$$\frac{dX}{dt} = f(X) - g(X) \quad (2.1)$$

Existen dos reacciones que afectan directamente la concentración de la proteína X : su síntesis y degradación. De acuerdo a la ley de acción de masas, la tasa de la reacción será proporcional a la concentración de los reactivos involucrados, de manera que las síntesis y degradación de X pueden ser representada por funciones dependientes de su propia concentración: $f(X)$ y $g(X)$, respectivamente. Aunque la concentración de X afecta directamente la expresión del mRNA y no de la proteína, asumimos que la concentración de esta última es proporcional a la de su mRNA, de manera que al modelar solo la concentración de la proteína estamos ignorando el retraso dado en el proceso de traducción. Por lo general, para $f(X)$ se utiliza la ecuación de Hill (2.2); donde K es la tasa de síntesis máxima de la proteína, K_D la constante de disociación de X con la región reguladora de su promotor y n es el coeficiente de Hill.

$$f(X) = K \frac{X^n}{K_D^n + X^n} \quad (2.2)$$

La utilidad de la ecuación de Hill radica en la forma de su curva (Figura 2.2); la forma sigmoideal y ultrasensible suele representar adecuadamente procesos resultantes de

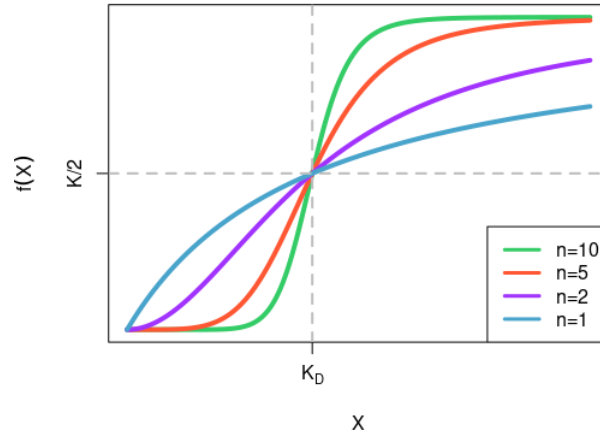


Figura 2.2: Función de Hill para regulación transcripcional positiva - Con parámetros $K = 80$, $K_D = 40$ y n con valores según se indica en la figura. El parámetro K_D es la concentración de X requerida para que la función de síntesis llegue a la mitad de su síntesis máxima.

varios pasos, como lo es el de transcripción. La ultrasensibilidad permite que cambios pequeños en la concentración de la proteína X produzcan cambios significativos en su síntesis; una manera de pensarlo es que si requieren de varias interacciones para que el proceso ocurra, este se vuelve más factible una vez que la proteína cruza cierto límite de concentración, pues dichas interacciones se vuelven más probables. En la Figura 2.2 se puede ver la forma de la función y el efecto de variar el coeficiente de Hill. Para la función con $n = 10$, cuando la concentración de X tiene valores pequeños el efecto que tiene sobre la activación de su síntesis es muy bajo, pero conforme se va acercando al valor de K_D la síntesis comienza a dispararse; por otra parte, al disminuir el coeficiente de Hill la fuerza del disparo va disminuyendo; al llegar a $n = 1$ donde la función ya no es sigmoideal.

Tradicionalmente, el coeficiente de Hill se ha asociado con el número de factores de transcripción que se unen de manera cooperativa entre ellos y el DNA para regular al gen correspondiente, pero el término de cooperatividad no es exclusivo a la formación de complejos de factores de transcripción. Por ejemplo, se ha visto que mecanismos post-traduccionales como la fosforilación de proteínas en múltiples sitios también puede representarse efectivamente con la ecuación de Hill (17).

La degradación de X está dada por la función $g(X)$. La degradación puede estar regulada por diversos mecanismos (e.g. marcas de degradación por el proceso de ubiquitina-

ción) que dependen de interacciones con otras proteínas o estímulos. Por simplicidad, aquí planteo un proceso de degradación por dilución donde la concentración de proteína disminuye gracias al crecimiento y división celular, haciendo de $g(X)$ una función lineal:

$$g(x) = \gamma X \quad (2.3)$$

Finalmente, la ecuación (2.1) se convierte en:

$$\frac{dX}{dt} = K \frac{X^n}{K_D^n + X^n} - \gamma X \quad (2.4)$$

En la sección 2.1.1 se mencionó que en ausencia de proteínas reguladoras el proceso de transcripción (y por consiguiente el de la traducción de los mRNAs) aún ocurre, aunque en menor proporción. La ecuación anterior no considera esto y si iniciáramos con una concentración de $X = 0$ no habría producción ni degradación de la proteína pues tanto $f(X)$ y $g(X)$ son igual a cero. En esta condición el sistema se encuentra en un estado estacionario, es decir $\frac{dX}{dt} = 0$. Un modelo más realista en el que se toma en cuenta la síntesis basal, siendo esta una pequeña proporción de la síntesis máxima representada por αK y donde $\alpha \in (0, 1)$, es el siguiente:

$$\frac{dX}{dt} = K(\alpha + (1 - \alpha) \frac{X^n}{K_D^n + X^n}) - \gamma X \quad (2.5)$$

Esta ecuación no tiene solución analítica; por lo general los sistemas de este tipo requieren de métodos numéricos para encontrar aproximaciones a la solución. Además de simulaciones computacionales, herramientas de dinámica no lineal suelen utilizarse para estudiar las dinámicas que puede generar el sistema (e.g. llegar a un estado estacionario); para ello se estudia la estabilidad de los puntos de equilibrio del sistema, también llamados estados estacionarios (i.e. cuando la tasa de cambio es igual a cero). En particular, dependiendo de los parámetros elegidos, el modelo de la ecuación (2.5) puede llegar a un estado estacionario único o presentar biestabilidad (16).

2.1.4. Biestabilidad

La retroalimentación positiva es capaz de dar lugar a la biestabilidad. La biestabilidad se caracteriza por contar con tres estados estacionarios: dos atractores (estables) y un repulsor (inestable). Aquí, dependiendo de las condiciones iniciales, las trayectorias del sistema se ven atraídas a uno de los dos estados estacionarios estables y repelidas del estado estacionario inestable, donde solo permanecen si la condición inicial corresponde a dicho punto (e.g. $X=0$ en la ecuación (2.4)). Además, este sistema tiene la capacidad de funcionar como un *switch* en el que señales externas al circuito pueden llevarlo de un estado estacionario a otro y se caracteriza por contar con un tipo de memoria llamada histéresis, en donde el nivel del estímulo necesario para *encender* el *switch* es distinto al nivel necesario para *apagarlo*, de manera que el estado en el que se encuentra la célula depende de su historia.

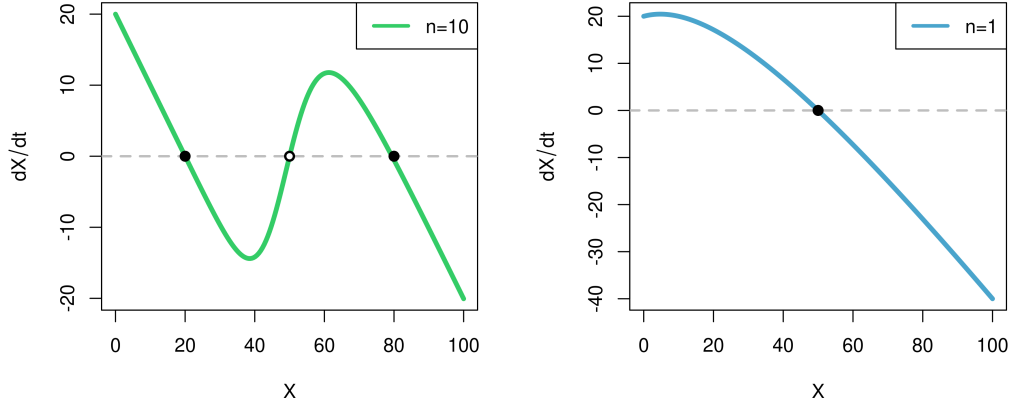


Figura 2.3: Biestabilidad - Tasa de cambio de X (ecuación 2.5) con respecto a X con parámetros $K = 80$, $\alpha = 0.25$, $K_D = 50$, $\gamma = 1$ y n según se indica en la figura. Con suficiente no linealidad ($n = 10$) el sistema es biestable mientras que con $n = 1$ el sistema es monostable (i.e. un solo punto de equilibrio). Los puntos negros representan puntos de equilibrio atractores y el punto blanco el punto de equilibrio repulsor.

En realidad, la biestabilidad es un caso de la multiestabilidad. Una manera útil de visualizar los estados estables es mediante pozos de energía potencial; cuando el estado del sistema se encuentra en un pozo, que corresponde a un estado estacionario estable, para cambiar de pozo debe cruzar una barrera de energía. Estos pozos pueden tener diferentes barreras energéticas, haciendo que de unos sea más o menos difícil salir, es decir que la fuerza de la perturbación para hacer el *switch* puede variar. En biología, estos pozos pueden verse como estados fenotípicos. Los circuitos que pueden producir multiestabilidad suelen estar presentes en procesos que requieren que la célula cambie de estado de manera irreversible (e.g. diferenciación) o de ser capaz de cambiar entre estados de manera dinámica (e.g. metabolismo de lactosa en *E. coli*) (36).

El modelo de la ecuación (2.5) representa el circuito génico más simple que puede presentar esta dinámica. En la izquierda de la Figura 2.3 se puede ver cómo para un conjunto de parámetros la tasa de cambio del factor de transcripción intersecciona tres veces con el eje x , dando lugar a los tres puntos de equilibrio mencionados. Si bien la retroalimentación positiva es necesaria para generar la dinámica, no es suficiente. Las tres intersecciones en este sistema son posibles gracias a que el valor de n es suficientemente alto para darle la forma sigmoideal de la ecuación de Hill; sin esta ultrasensibilidad, es decir si $n = 1$, solo se obtiene un estado estacionario estable (Figura 2.3, derecha).

2.1.5. Oscilaciones

Las dinámicas oscilatorias están presentes en varios procesos de las células, dos de los más comunes son los ciclos circadianos y celulares. Los ciclos circadianos permiten a diversos organismos expresar ciertas concentraciones de moléculas en ciclos con periodo de 24 horas. En cuanto al ciclo celular, mientras algunos organismos optan por realizar su división como una serie de pasos en cadena, otros como el anfibio *Xenopus laevis* lo hace de manera cíclica. En ambos casos, el mecanismo detrás es un reloj autónomo que funciona gracias a un circuito génico con retroalimentación negativa.

La retroalimentación negativa es necesaria para generar oscilaciones; aunque solo ciertos circuitos cumplen con las condiciones requeridas para oscilar. Por ejemplo, un circuito de un solo elemento que se autoregula de manera negativa (modelado con la ecuación (2.5) pero con valores negativos para el coeficiente de Hill, haciendo al TF un represor en lugar de un activador), no puede presentar oscilaciones y siempre llega a un estado estacionario estable. Para que un circuito exhiba oscilaciones debe existir un retraso temporal que impida al sistema alcanzar su estado estacionario. Existen dos maneras en las que este puede generarse: incorporando el retraso al circuito de retroalimentación negativa, ya sea introduciendo más elementos o utilizando ecuaciones diferenciales con retraso temporal explícito, o añadiendo retroalimentación positiva (28), (11).

Un modelo clásico de oscilador biológico es el de Goodwin (20) (Figura 2.4). Su sistema consta de tres variables que corresponden al mRNA (X), a su proteína (Y) y a un inhibidor transcripcional (Z) que es activado por Y y que actúa sobre X; está dado por el siguiente sistema de ecuaciones diferenciales:

$$\frac{dX}{dt} = K_1 \frac{K_D^n}{K_D^n + Z^n} - K_2 X \quad (2.6)$$

$$\frac{dY}{dt} = K_3 X - K_4 Y \quad (2.7)$$

$$\frac{dZ}{dt} = K_5 Y - K_6 Z \quad (2.8)$$

Solo para algunos parámetros el modelo presenta el comportamiento requerido: un ciclo límite (i.e. órbitas cerradas a donde se dirigirán las trayectorias del sistema (32)). En particular, la no-linealidad juega un papel importante; para el caso del modelo de Goodwin el valor mínimo para general oscilaciones es $n = 8$ (17). Como se mencionó antes, añadir elementos (el mRNA y una proteína extra en este caso) a la retroalimentación negativa no es la única manera de generar retraso. Al modelar la retroalimentación negativa en células eucariontes, el retraso temporal explícito en las ecuaciones diferenciales puede venir de la exportación del mRNA al citoplasma y posteriormente la importación de la proteína reguladora al núcleo.

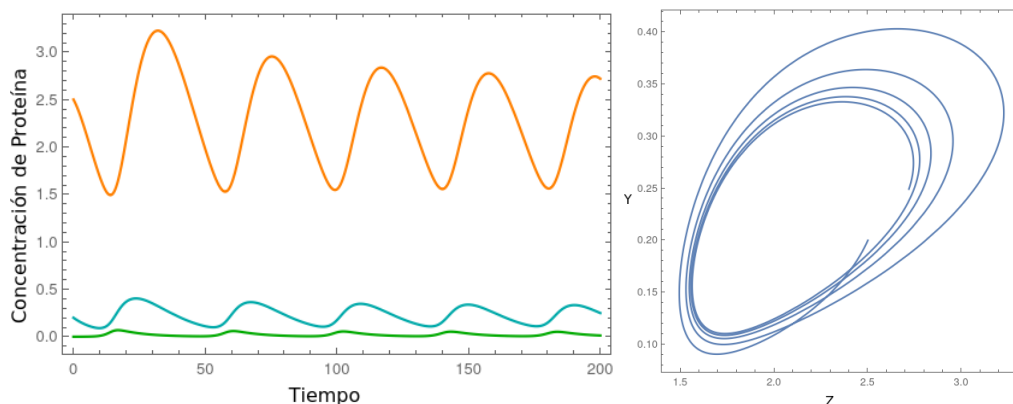


Figura 2.4: Oscilador de Goodwin - Simulación determinista del oscilador de Goodwin con parámetros: $K_1 = K_D = K_3 = K_5 = 1$, $K_2 = K_4 = 0.1$, $K_6 = 0.11$ y $n = 10$; con condiciones iniciales $X = 0$, $Y = 0.2$ y $Z = 2.5$. En la izquierda se muestran las concentraciones de las proteína y mRNA con respecto al tiempo. En la derecha el ciclo límite.

La retroalimentación positiva también puede evitar que un sistema alcance su estado estacionario. Las oscilaciones de este tipo de circuito son conocidas como *oscilaciones de relajación* y son cualitativamente distintas a las del otro tipo de oscilador (e.g. modelo de Goodwin), pues tienen una fase rápida y una lenta que les da un aspecto menos suave (11). Además, se ha observado que la retroalimentación positiva hace que las oscilaciones mantengan una amplitud robusta ante cambios en los parámetros, pero a costo de que el periodo se vuelva más sensible (21). Estas diferencias pudieran tener roles funcionales haciendo a ciertos circuitos ventajosos bajo ciertas condiciones. Los osciladores con retroalimentación positiva no han sido observados en ciclos circadianos robustos, pero sí han sido ampliamente utilizados en modelos del ciclo celular (11).

2.2. Ruido Bioquímico

La expresión génica es intrínsecamente estocástica. Desde el pegado de un factor de transcripción hasta la producción final de la proteína hay una serie de eventos probabilísticos que involucran interacciones de moléculas discretas, causando que la expresión génica no sea uniforme aún en poblaciones celulares isogénicas. Desde hace varios años se ha observado que el nivel de expresión génica es variable entre células, pero solo recientemente y gracias a que estas dinámicas se han logrado estudiar a nivel de células individuales, el fenómeno y sus causas han sido caracterizados a más detalle (29). Las causas pueden clasificarse como extrínsecas e intrínsecas. Las primeras se refieren a las variaciones que pueden ocurrir entre células, como el número de ribosomas o RNA

polimerasas. Las segundas son variaciones inherentes al proceso de transcripción, por lo que aún causarían fluctuaciones en una población genéticamente idéntica y con concentraciones iguales de componentes celulares (9).

La producción de proteínas se da en *disparos* estocásticos y no de manera continua y uniforme. En Cai et al. (6), a través de la cuantificación a lo largo del tiempo de la enzima β -galactosidasa en células individuales, se encontró que el número de moléculas era constante durante algunos periodos pero había momentos en los que incrementaba, de manera que en cada *disparo* muchas moléculas eran sintetizadas y después venían periodos donde no había producción. Las fuentes de estos disparos vienen principalmente del proceso de transcripción. En Golding et al. (14), al contabilizar el número de moléculas de mRNA en *E. coli* sobre una serie de tiempo, observaron que la transcripción consta de periodos largos donde no hay actividad y de periodos intensos y cortos donde se sintetizan los transcritos. En células eucariontes en particular, estos estados de *encendido* y *apagado* también se han atribuido a rearrreglos en la posición de los nucleosomas sobre los promotores (29). En general, el hecho de que en una célula existan solo una o dos copias por gen, hace que las fluctuaciones tengan efectos notables. Cuando el promotor se encuentra en el estado de encendido y se produce el disparo de transcritos, este disparo es amplificado cuando dichos transcritos se traducen, haciendo que las proteínas también se sinteticen en disparos.

Otros aspectos importantes del ruido bioquímico son el tiempo y la propagación. El primero se refiere a que la acumulación de proteínas a lo largo del tiempo puede amortiguar la variabilidad generada por la expresión en *disparos*. El segundo a que la fluctuaciones pueden ser transmitidas a lo largo de los componentes de las redes de regulación, de manera que la variabilidad de genes en vías superiores puede añadir a la de el gen o los genes que regula (7).

2.2.1. Modelos estocásticos

En la sección 2.1.3 se modeló la expresión génica considerando la concentración de la proteína y utilizando ecuaciones diferenciales, pero al modelar el sistema de esta forma se ignora el efecto del ruido bioquímico y la naturaleza discreta de las proteínas. En muchos sistemas con grandes cantidades de moléculas el efecto estocástico puede ignorarse, pues se promediará y el ruido podría no tener implicaciones importantes, pero este no es el caso en la mayoría de los sistemas biológicos in vivo. Esencialmente, los modelos estocástico toman en cuenta que las reacciones son procesos aleatorios que pueden ocurrir con cierta probabilidad, llamada *propensidad* (35).

Gillespie desarrolló una técnica llamada *Stochastic Simulation Algorithm (SSA)*, conocida como el algoritmo de Gillespie, que implementa un procedimiento exacto para simular las dinámicas químicas estocásticas de los sistemas bioquímicos (13). Aquí, el nivel de expresión génica (e.g. el número de proteínas de las especies involucradas en

las reacciones) es considerado como un estado del sistema y las dinámicas de expresión como una caminata aleatoria sobre los posibles estados. El sistema cambia de estado cuando una reacción ocurre (e.g. degradación de una molécula); el tiempo entre cada reacción decrece de manera exponencial y su *propensidad* depende del estado actual del sistema; es decir que es proporcional al número de moléculas involucradas en la reacción. Por ejemplo, para el modelo de la ecuación (2.5) la propensidad de la síntesis de la proteína X corresponde a $K(\alpha + (1 - \alpha)\frac{X^n}{K_D^n + X^n})$ y la propensidad de su degradación a γX .

El algoritmo funciona de la siguiente manera. Dado el estado del sistema (i.e. el número de elementos de las especies) se calculan las propensidades de las reacciones. Después se elige una de las reacciones, de manera que la elección de una reacción específica es proporcional a su propensidad, y se eligen el tiempo en el que ocurrirá la reacción, el cual se toma de una distribución exponencial. Finalmente, el estado del sistema y el tiempo se actualizan de acuerdo a la reacción que ocurrió. Estos pasos se realizan hasta alcanzar el tiempo establecido para la simulación.

2.2.2. *State-switching*

Tradicionalmente se ha visto al ruido bioquímico como un obstáculo con el que la célula debe lidiar para llevar a cabo sus funciones; por ejemplo, se ha especulado que la presencia de motivos de retroalimentación negativa podría estar relacionada con su función de reducir la variación en la expresión génica (1). Pero existen varios casos en los que el ruido juega un papel funcional; la *diferenciación probabilística* es uno de ellos. Esta permite a células genéticamente idénticas alcanzar destinos fenotípicos distintos sin la necesidad de señales externas o de desarrollar complejos mecanismos de toma de decisión (e.g. cuando existen muchos destinos fenotípicos una elección estocástica puede llegar a ser más sencilla) (7) (25). Aunque existen diversos mecanismos de diferenciación probabilística, esta sección se centra en el de *state-switching*.

El *state-switching* o *switching* epigenético requiere de biestabilidad (o multiestabilidad). Las células que se encuentren en un determinado estado estacionario pueden ser empujadas fuera de él gracias a las fluctuaciones en la concentración de sus proteínas provocadas por el ruido bioquímico; retomando la analogía de los pozos de potencial, el ruido puede llevar a la célula de un pozo a otro. Simulaciones estocásticas con el algoritmo de Gillespie del modelo de la ecuación (2.5) resultan en una distribución bimodal de la concentración de la proteína, donde las modas corresponden a los estados estacionarios estables indicados en la Figura 2.5. La distribución también es indicativa de la probabilidad de transición entre los estados, siendo la moda mayor el estado estacionario del que es más difícil salir. La frecuencia de transición depende de los parámetros y como se verá en secciones posteriores, ciertas frecuencias pueden resultar benéficas en algunos contextos.

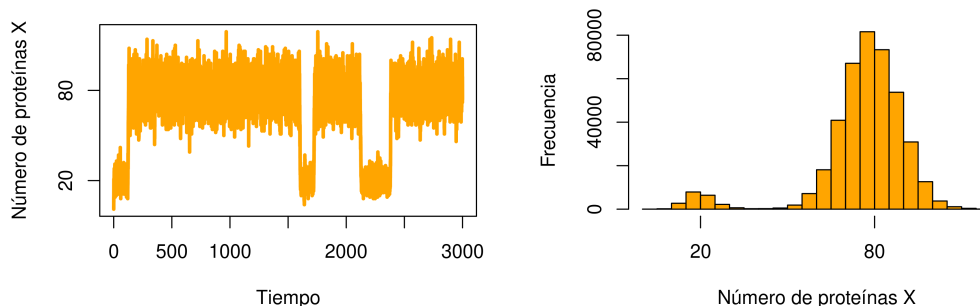


Figura 2.5: *Switch* Epigenético - Simulación estocástica del modelo descrito en la sección 2.1.3, ecuación (2.5). En la izquierda se muestra la serie de tiempo de la simulación del número de moléculas de la proteína X ; a la derecha su histograma. Las modas de la distribución corresponden a los estados estacionarios del sistema: $X = 20$ y $X = 80$

2.2.3. Oscilaciones ante el ruido bioquímico

Los osciladores bioquímicos son capaces de generar dinámicas cíclicas precisas y robustas a las fluctuaciones que enfrentan dentro de las células (e.g. ruido bioquímico). Por ello se ha especulado que las redes de regulación pudieran tener restricciones topológicas que le confieran al circuito robustez; dentro de ellas se ha propuesto a la retroalimentación positiva (2), aunque esto sigue siendo controversial. Recientemente, en Li et al. (24) se analizó la robustez (cuantificada como el espacio de parámetros que presenta oscilaciones) de todos los circuitos de tres nodos con topologías oscilatorias. Se encontró que las estructuras centrales de la red, es decir los circuitos que no podían simplificarse más sin perder la capacidad de presentar la dinámica oscilatoria, determinan el nivel de robustez básico. Los circuitos centrales más robustos que identificaron concuerdan con las estructuras más comunes en circuitos biológicos y sintéticos, siendo estas el *repressilator* (8), el activador-represor (i.e. circuito con retroalimentación positiva) y la retroalimentación negativa con retraso (e.g. modelo de Goodwin), esquematizados en la Figura 2.6. Además, encontraron que había un efecto positivo en la robustez al añadir nodos con *input* incoherente (i.e. un nodo recibe activación e inhibición); sugiriendo que ciertas estructuras pudieran haber sido seleccionadas por su capacidad de generar osciladores robustos.

Por otra parte, varios trabajos han reportado un rol *constructivo* del ruido en las dinámicas de modelos de osciladores bioquímicos. En modelos de relojes circadianos (i.e. osciladores bioquímicos con periodos de 24 horas), cerca de puntos de bifurcación Hopf (i.e. región en la que un cambio en los parámetros causa un cambio cualitativo en la estabilidad del sistema, surgiendo la solución periódica necesaria para generar

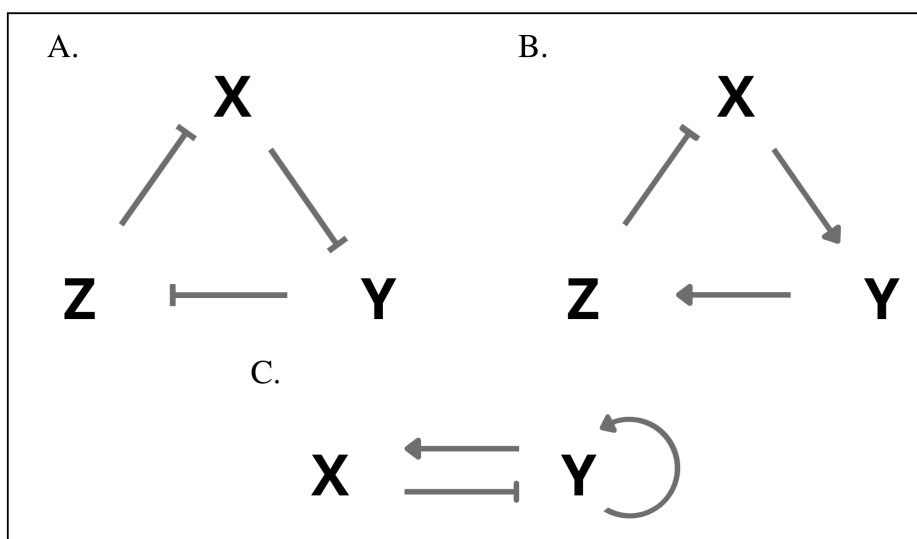


Figura 2.6: Circuitos centrales de osciladores robustos - Las tres estructuras más robustas ante perturbaciones en el espacio de parámetros. **A)** *Repressilator*. **B)** Retroalimentación negativa con retraso. **C)** Activador-represor.

oscilaciones) se han observado oscilaciones en los modelos estocásticos, mientras que en los deterministas el sistema llega a un estado estacionario, pues su estabilidad es la de un punto de equilibrio atractor y no la de un ciclo límite (18), (19), (38). Estas oscilaciones tienen periodo variable y amplitudes altas, pero al irse acercando al punto de bifurcación se asemejan más a las que se obtienen en el sistema determinista cuando este presenta ya el ciclo límite. La dinámica que proponen es la siguiente: el estado estacionario estable es excitable, de manera que perturbaciones pequeñas lleven al sistema fuera de él y generan recorridos largos por el plano fase, correspondientes a un *disparo* en la transcripción, generando oscilaciones irregulares. Además, se ha observado que el ruido puede aumentar las amplitudes de oscilaciones que en simulaciones deterministas presentan baja amplitud.

Se ha especulado que los organismos podrían tomar ventaja de las oscilaciones inducidas por ruido. En Li et al. (23) observaron que mientras que el modelo determinista del reloj circadiano de *Drosophila* existe una región precisa de oscilaciones en el espacio de parámetros, el modelo estocástico amplía esta región conservando periodos similares; proponen que esta región puede proporcionarle robustez al oscilador, de forma que ante cambios en los parámetros (que en el modelo determinista llevan al sistema a un estado estacionario) el reloj permanezca funcional.

Hasta ahora y bajo mi conocimiento, solo se ha reportado un caso biológico en el que el papel del ruido resulta fundamental en ciertos contextos para el funcionamiento de un

oscilador bioquímico. En Ko et al. (22), a través de experimentos y modelaje matemático observaron que las células del núcleo supraquiasmático presenta ritmos aún cuando su reloj circadiano es disfuncional debido a mutaciones en los componentes del circuito. Ellos proponen que a través del ruido bioquímico y de interacciones intercelulares las células logran generar oscilaciones.

2.3. Adaptación ante ambientes fluctuantes

Los organismos se encuentran en ambientes dinámicos que pueden influir directamente en aspectos fundamentales para su supervivencia. Los cambios ambientales a los que se enfrentan pueden ser cíclicos, como la temporada estacional y el ciclo diurno, o repentinos, como la escasez de recursos por perturbaciones externas. Para lidiar con ellos, un amplio repertorio de estrategias han sido desarrolladas; actualmente, bajo qué condiciones evolutivas surgen y se seleccionan estas estrategias permanece siendo una amplia e importante pregunta en la biología evolutiva.

En general, los organismos pueden implementar dos tipos de respuesta: mecanismos robustos ante las fluctuaciones ambientales que les permitan mantener su estado fenotípico (i.e. homeostasis) o generando variación fenotípica que les permita ajustarse al futuro estado ambiental (i.e. variación adaptativa) (27). Este capítulo se centra en la segunda categoría y en algunas de las estrategias que derivan de ella. Aunque existen varias maneras en las que estas vertientes se podrían clasificar (27), (15), (4), por simplicidad considero dos grupos: variación genética (i.e. a través de mutaciones) y no genética (i.e. a partir de un mismo genoma, por ejemplo cuando los organismos aprovechan las propiedades emergentes de sus circuitos de regulación para generar variación en su expresión génica).

2.3.1. Variación genética

Experimentos clásicos han demostrado que a través de mutaciones las poblaciones generan diversidad fenotípica y gracias a ello, cuando son sometidas ante un cambio ambiental que produce presión selectiva, los individuos con el fenotipo que confiera ventaja adaptativa serán seleccionados (26); a esta estrategia se le llamada adaptación genética. Bajo condiciones de estrés, algunos organismos han desarrollado mecanismos que les permiten aumentar su variación genética. Por ejemplo, las bacterias tienen diversas maneras de aumentar su tasa de mutación; una de ellas es la respuesta SOS. Ésta normalmente es inducida ante daño en el DNA, pero se ha observado que, al menos parcialmente, también puede ser activada en ambientes de estrés. La respuesta SOS consiste en la activación de varios genes, dentro de los cuales están DNA polimerasas pensadas a errores (12).

2.3.2. Variación no genética

Plasticidad fenotípica

La plasticidad fenotípica se refiere a la capacidad de un organismo de cambiar su fenotipo de acuerdo a señales ambientales. Los relojes circadianos son un ejemplo; aunque fundamentalmente son osciladores bioquímicos autónomos, pues en ausencia de las señales ambientales (i.e. luz) continúan oscilando con periodos precisos que se conservan por varias generaciones, pueden adaptarse a los periodos del ciclo diurno de su ambiente. Estos relojes son implementados por organismos tanto unicelulares como pluricelulares para regular procesos como expresión génica, actividad locomotora, niveles hormonales, absorción de nutrientes y otros (3). En particular, las plantas los utilizan para regular positivamente la maquinaria de fotosíntesis antes del amanecer de manera que esté lista para responder a la luz solar (4).

Bet-hedging

Esta estrategia ocurre cuando la población tiene fenotipos que presentan desventajas adaptativas ante un ambiente, pero que ante otro confieren una ventaja. La resistencia a antibióticos por poblaciones de *E. coli* es un ejemplo (25). Se ha observado que una pequeña proporción de la población, llamadas células persistentes, se encuentran en un estado de nulo o de bajo crecimiento que les permite evadir el efecto de los antibióticos. De esta forma las células que se encontraban en división celular mueren, pero las persistentes sobreviven y reanudan su crecimiento. El *state switching* o *switching* epigenético también es una forma de *bet-hedging* en la que a través de las fluctuaciones generadas por ruido, como mencionó en la sección 2.2.2, el sistema puede transicionar de un estado a otro; en este caso, uno de los estados fenotípicos (i.e. nivel de expresión génica) resulta ventajoso en alguno de los ambientes.

2.3.3. Condiciones ambientales que favorecen a las estrategias adaptativas

La naturaleza de las fluctuaciones ambientales es un factor importante que influye en que estrategia es más adecuada; uno de sus aspectos es la predictibilidad. Si se está lidiando con un cambio periódico, es posible que mecanismos que puedan anticipar el cambio sean benéficos. Por otra parte, si los organismos se enfrentan ante cambios aleatorios o a cambios para los que no se pueda percibir un patrón claro (e.g. si la frecuencia de la fluctuación sobrepasa el tiempo de vida del organismo), la selección podría favorecer estrategias de *bet-hedging* (27). Además, para desarrollar mecanismos de plasticidad fenotípica, los organismos requieren de señales ambientales confiables que no siempre están a su disposición, en cuyo caso una estrategia de *bet-hedging* podría ser superior.

2. MARCO TEÓRICO

Trabajos experimentales y teóricos han demostrado que el *bet-hedging* puede ser óptimo cuando la tasa de cambio entre estados está sincronizada con el cambio típico ambiental y cuando la presión selectiva es simétrica entre los ambientes (e.g. (30) y (34)). En Gómez-Schiavon et al. (16), a través de simulaciones computacionales, evaluaron como competían las estrategias de adaptación genética y *switching* epigenético. Encontraron que el *bet-hedging* es superior a la adaptación genética cuando el ambiente fluctúa a tasas altas, lo que podría estar relacionado con que la frecuencia a la que ocurren las *epimutaciones* (i.e. las transiciones entre estados fenotípicos) es mayor a la de las mutaciones en la adaptación genética.

3.1. Desarrollo del modelo

Con el fin de estudiar las dinámicas de un circuito génico capaz de presentar oscilaciones y biestabilidad, se desarrolló un modelo de ecuaciones diferenciales. Para presentar ambas dinámicas, el circuito modelado debe contar con retroalimentación positiva y negativa; pues como se mencionó en el Capítulo 2, la retroalimentación positiva es necesaria para la biestabilidad, mientras que la retroalimentación negativa para las oscilaciones. Para mantener el número de dimensiones mínimo y facilitar la caracterización del modelo, el circuito consiste solo de dos genes: A, quien puede actuar como un activador y B, quien puede actuar como un represor. La regulación de estos genes se da a nivel transcripcional de la siguiente manera: la proteína A es un factor de transcripción capaz de activar su propia síntesis (retroalimentación positiva) así como la de la proteína B, quien también es un factor de transcripción capaz de reprimir la síntesis de A (retroalimentación negativa). De esta manera tanto A como B compiten por el promotor de A. En la Figura 3.1 se esquematizan estas interacciones.

Ya que ambas proteínas regulan la síntesis de A, esta es una función de A y B; mientras que la síntesis de B solo es regulada por A, siendo una función de A. Por simplicidad asumimos que la degradación de las proteínas solo es por dilución, de tal manera que sus funciones de degradación solo dependen de ellas mismas; por lo tanto, el modelo de ecuaciones diferenciales del sistema tiene la siguiente forma:

$$\frac{dA}{dt} = f_A(A, B) - g_A(A) \quad (3.1)$$

$$\frac{dB}{dt} = f_B(A) - g_B(B) \quad (3.2)$$

En nuestro modelo, la función de síntesis de B (f_B) sigue una ecuación Hill de activación simple que considera síntesis basal, es decir que puede ser sintetizada en ausencia de A a una pequeña proporción de su síntesis máxima de acuerdo a un parámetro $\beta \in [0, 1)$.

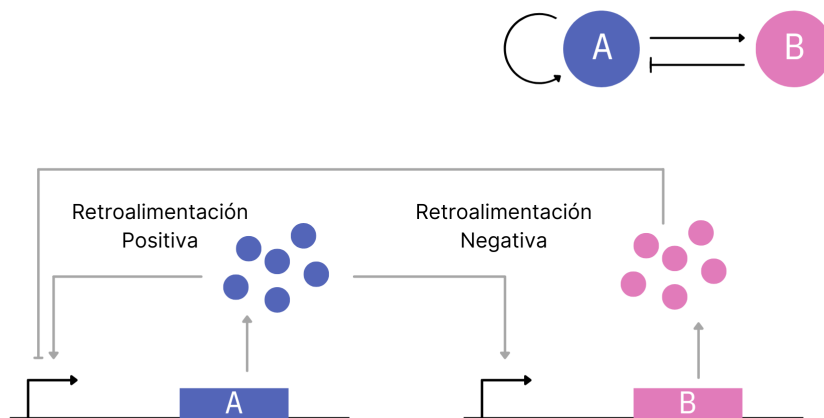


Figura 3.1: Diagrama del circuito de retroalimentación positiva y negativa - El gen A se auto-activa al mismo tiempo que activa al gen B, quien reprime a A.

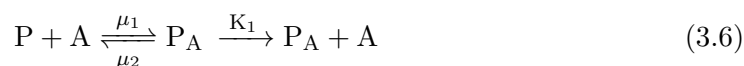
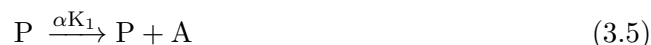
En cuanto a las funciones de degradación, éstas serán lineales para representar pérdida simple por dilución. Con estas consideraciones, las ecuaciones diferenciales quedan de la siguiente forma:

$$\frac{dA}{dt} = f_A(A, B) - K_2A \quad (3.3)$$

$$\frac{dB}{dt} = K_3\left(\beta + (1 - \beta)\frac{A^n}{A^n + K^n}\right) - K_4B \quad (3.4)$$

Donde K_2 es la tasa de dilución de A, K_3 la tasa de síntesis máxima de B, β es la actividad basal relativa a K_3 , K es la constante de disociación de B con el promotor de A, n es el coeficiente de Hill asociado a la regulación de B sobre el promotor de A y K_4 la tasa de dilución de B.

Para obtener una función f_A que describa adecuadamente la competencia de ambas proteínas por el promotor de A, consideramos primero las ecuaciones cinéticas del proceso:



La síntesis basal está representada en la ecuación (3.5), en ella la síntesis se puede llevar acabo con el promotor libre (P) a una tasa proporcional a la de síntesis máxima: αK_1 , donde K_1 representa la síntesis máxima y α la actividad basal relativa en donde $\alpha \in [0, 1)$. La unión de A a su promotor y su síntesis por activación están representadas en la ecuación (3.6), donde la proteína A puede unirse y desunirse al promotor con tasas μ_1 y μ_2 , respectivamente; cuando la proteína está unida al promotor (P_A), la tasa de síntesis corresponde a K_1 . Finalmente, la ecuación (3.7) describe la unión de la proteína B al promotor de A, impidiendo así la síntesis de A, tanto por la actividad basal del promotor libre P, o mediante la activación de A. B se asocia y disocia del promotor a una tasa μ_3 y μ_4 , respectivamente.

De acuerdo a lo anterior, el promotor puede estar en tres estados: libre (P), unido a A (P_A) o unido a B (P_B); por lo tanto, el total de promotor está representado por la suma de los tres estados, que sin pérdida de generalidad, asumimos es igual a 1:

$$P_T = P_A + P_B + P \quad (3.8)$$

$$P_T = 1 \quad (3.9)$$

$$P_A + P_B + P = 1 \quad (3.10)$$

Entonces, asumiendo los pasos cinéticos descritos en las ecuaciones (3.5)-(3.7) siguen el principio de acción de masas, las funciones de cambio de los tres estados del promotor son las siguientes:

$$\frac{dP}{dt} = \mu_2 P_A + \mu_4 P_B - \mu_1 P A - \mu_3 P B \quad (3.11)$$

$$\frac{dP_A}{dt} = \mu_1 P A - \mu_2 P_A \quad (3.12)$$

$$\frac{dP_B}{dt} = \mu_3 P B - \mu_4 P_B \quad (3.13)$$

Como la síntesis de A se lleva acabo con el promotor en sus estados libre y unido a A, su función de síntesis puede reescribirse como:

$$f_A(A, B) = K_1(\alpha(P) + P_A) \quad (3.14)$$

Despejando P de la ecuación (3.10) obtenemos:

$$P = 1 - P_A - P_B \quad (3.15)$$

Y esta ecuación se sustituye en la función (3.14):

$$f_A(A, B) = K_1(\alpha(1 - P_A - P_B) + P_A) \quad (3.16)$$

3. METODOLOGÍA

Suponiendo que las interacciones entre las proteínas A y B con el promotor de A ocurren de manera más rápida que el resto del proceso de transcripción, se puede realizar una aproximación de *quasi-equilibrio*. En ella se asume que las tasas de cambio de P_A y P_B alcanzan equilibrio antes que el resto del sistema:

$$\frac{dP_A}{dt} = 0 \quad (3.17)$$

$$\frac{dP_B}{dt} = 0 \quad (3.18)$$

El promotor libre P es sustituido por la ecuación (3.15) en las ecuaciones diferenciales de P_A (3.12) y P_B (3.13), de ellas se obtienen los siguientes valores:

$$P_A = \frac{A}{A + \frac{\mu_2\mu_3}{\mu_1\mu_4}B + \frac{\mu_2}{\mu_1}} \quad (3.19)$$

$$P_B = \frac{B}{B + \frac{\mu_4}{\mu_3}}(1 - P_A) \quad (3.20)$$

Definiendo las constantes de disociación como $K_A = \frac{\mu_2}{\mu_1}$ y $K_B = \frac{\mu_4}{\mu_3}$, se obtiene:

$$P_A = \frac{A}{A + \frac{K_A}{K_B}B + K_A} \quad (3.21)$$

$$P_B = \frac{B}{B + K_B}(1 - P_A) \quad (3.22)$$

Finalmente, se sustituyeron los valores de P_A y P_B de las ecuaciones (3.21) y (3.22) en la función de síntesis (3.16) para obtener la versión final:

$$f_A(A, B) = K_1 \left(\alpha \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} + (1 - \alpha) \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} \right) \frac{A^{n_A}}{A^{n_A} + \frac{K_A^{n_A}}{K_B^{n_B}}(B^{n_B} + K_B^{n_B})} \quad (3.23)$$

Para darle la no-linealidad requerida para las dinámicas de interés, se incluyen coeficientes de Hill de tal manera que en ausencia de A el modelo representa un ecuación de represión de Hill simple, en donde B actúa como represor de A, y en ausencia de B una ecuación de activación de Hill simple, en donde A actúa como su propio activador. El modelo está dado por las siguientes ecuaciones diferenciales:

$$\frac{dA}{dt} = K_1 \left(\alpha \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} + (1 - \alpha) \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} \right) \frac{A^{n_A}}{A^{n_A} + \frac{K_A^{n_A}}{K_B^{n_B}}(B^{n_B} + K_B^{n_B})} - K_2 A \quad (3.24)$$

$$\frac{dB}{dt} = K_3(\beta + (1 - \beta) \frac{A^n}{A^n + K^n}) - K_4 B \quad (3.25)$$

Como se mencionó en el Capítulo 2, los coeficientes de Hill pueden tener diferentes interpretaciones y para no perder generalidad en el modelo no se considera ninguna en particular.

3.1.1. Simulaciones Deterministas del sistema

Para las simulaciones mostradas en las figuras de los Capítulos 2, 4 y 5 se utilizó la función *DSolve* de Wolfram Mathematica; con los parámetros y condiciones iniciales que se indican en las figuras.

3.1.2. Análisis de estabilidad

El análisis de estabilidad permite caracterizar cualitativamente la solución de un sistema de ecuaciones diferenciales ordinarias sin tener que resolverlo. Esto lo logra encontrando los estados solución del sistema (i.e. a dónde se dirigen las trayectorias) a través de estudiar sus puntos de equilibrio. Basándome en la teoría de dinámica no lineal presente en Strogatz (32), dado un conjunto de parámetros, el procedimiento empleado para realizar el análisis de estabilidad se describe en esta sección.

En el análisis consideramos dos pasos: primero, encontrar los estados estacionarios del sistema y segundo, evaluar su estabilidad. Sistemas no lineales como el presente suelen linealizarse sobre sus puntos de equilibrio para así evaluar su estabilidad de acuerdo a la clasificación de los sistemas lineales. Esta técnica parte del supuesto de que el sistema no lineal se comporta como uno lineal dentro de la vecindad de su punto de equilibrio. Para encontrar la aproximación a dicho punto, se utiliza la serie de Taylor sobre el punto de equilibrio hasta su término lineal, resultando en la matriz jacobiana.

El sistema se encuentra en estado estacionario cuando todas sus ecuaciones diferenciales son igual a cero. Esto ocurre en el o los puntos de intersección de $\frac{dA}{dt} = 0$ y $\frac{dB}{dt} = 0$. A la curva $\frac{dA}{dt} = 0$ se le conoce como *nullcline* de A y a la curva $\frac{dB}{dt} = 0$ como *nullcline* de B .

La *nullcline* de A está dada por:

$$h_A(A) = \left(\frac{A^{n_A} K_1 K_B^{n_B} - A^{1+n_A} K_B^{n_B} + \alpha K_1 K_A^{n_A} K_B^{n_B}}{A K_A^{n_A} - K_B^{n_B}} \right)^{\frac{1}{n_B}} \quad (3.26)$$

La *nullcline* de B por:

$$h_B(A) = \frac{A^n K_3 + \beta K_3 K^n}{A^n + K^n} \quad (3.27)$$

El estado estacionario se obtiene al resolver:

$$h_A(A) = h_B(A) \quad (3.28)$$

Un ejemplo se encuentra en la Figura 3.2. La ecuación (3.28) no tiene solución analítica. Para obtener una aproximación de la intersección de las *nullclines* y realizar el análisis de estabilidad desarrollé un código en MATLAB. En él se evalúan las *nullclines* en el rango de valores $(0, K_1]$ (se escogió ese rango de valores porque como se puede ver en la Figura 3.2, la *nullcline* de A empieza a decrecer cuando se acerca a K_1) y se resta

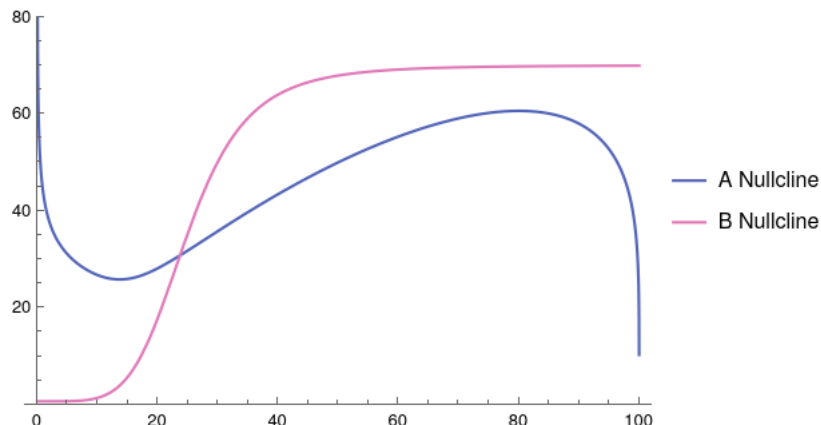


Figura 3.2: Nullclines - Parámetros de la *nullcline* de A: $K_1 = 100$, $\alpha = 0.5$, $K_A = 20$, $K_B = 20$, $n_A = 5$, $n_B = 5$, $K_2 = 1$. Parámetros de la *nullcline* de B: $K_3 = 70$, $\beta = 0.01$, $K = 25$, $n = 5$, $K_4 = 1$

la *nullcline* h_B (3.27) de la *nullcline* h_A (3.26), cuando se detectan cambios de signo (indicativos de que ocurrió una intersección) se evalúa más extensivamente la región y se determinan los estados estacionarios. La o las soluciones resultan en valores de A en su estado estacionario A_{ss} . Cada uno de ellos es evaluado en alguna de las *nullclines* para obtener el componente B_{ss} . Posteriormente, los puntos de equilibrio (A_{ss}, B_{ss}) se evalúan en la matriz jacobiana, y de esta se obtienen los valores propios. El código no asigna directamente un comportamiento, sino que contabiliza las intersecciones encontradas, la estabilidad de sus estados estacionarios (estables o inestables), si los eigenvalores tienen o no componente compleja y si las *nullclines* se vuelven discontinuas. A partir del resultado del código, y de acuerdo a la clasificación de sistemas lineales, se le asigna un comportamiento. Si llegan a presentar discontinuidad el resultado obtenido debe analizarse con más cuidado. El código se encuentra en Apéndice 1 (A.1)

3.1.3. Exploración del espacio de parámetros

Para caracterizar el espacio de parámetros se tomaron dos caminos. El primero fue utilizando el método *System Design Space*, en donde el sistema se divide en fenotipos con soluciones y propiedades distintas; el segundo consiste en tomar varias combinaciones de parámetros por medio de un muestro Monte Carlo y realizar análisis de estabilidad de su o sus estados estacionarios.

3.1.3.1. Exploración basada en fenotipos

System Design Space (10, 31) es un método creado para explorar el potencial fenotípico de sistemas bioquímicos; en esencia facilita la caracterización de los posibles

estados estacionarios del sistema desde una perspectiva fenotípica. En lugar de determinar el comportamiento de ciertos conjuntos de parámetros realizando un análisis de estabilidad (como el que se hizo en la sección 3.1.2), el programa define fenotipos con ciertas características a los que después les asigna parámetros representativos. Los fenotipos están definidos por soluciones cualitativamente distintas sobre su estado estacionario, proporcionando así regiones en el espacio con cierta estabilidad. Su ventaja radica en que se puede explorar el espacio rápidamente de forma global; como el modelo desarrollado en este trabajo tiene 10 parámetros libres, evaluar la estabilidad de combinaciones de ellos resulta limitado y computacionalmente costoso.

Para aplicar este método, el modelo primero debe reescribirse en la forma de *Generalized Mass Action* (GMA). Esta representación permite escribir la tasa de cambio de cada especie (e.g. proteínas) como la suma de los productos positivos y negativos del sistema:

$$\frac{dX_i}{dt} = \sum_{k=1}^{P_i} \alpha_{ik} \prod_{j=1}^Q X_j^{g_{ijk}} - \sum_{k=1}^{N_i} \beta_{ik} \prod_{j=1}^Q X_j^{h_{ijk}} \quad i = 1 \dots M \quad (3.29)$$

Donde X_i representa la concentración de una especie cuyo cambio se describe con una ecuación diferencial ordinaria y no lineal; en total hay M ecuaciones y Q concentraciones. Para los términos positivos los parámetros (e.g. tasas de síntesis) son α_{ik} y los exponentes son g_{ijk} ; para los términos negativos los parámetros y exponentes corresponden a β_{ik} y a h_{ijk} , respectivamente. P_i corresponde al número de términos positivos y N_i al de términos negativos de la ecuación i .

Para encontrar los distintos fenotipos, es decir las soluciones del sistema a sus posibles estados estacionarios, se hace una aproximación llamada *Dominant S-system*. En ella se asume que uno de los términos positivos y uno de los términos negativos dominan sobre los demás en cada ecuación; es decir, dos reacciones dominan a las otras. Esto resulta en la construcción de subsistemas (*S-systems*) en los que el sistema original se reduce a un solo término positivo y negativo para el que se busca un estado estacionario al igualarlo a cero. Esto se representa como:

$$\frac{dX_i}{dt} = \alpha_{ip} \prod_{j=1}^Q X_j^{g_{ijp}} - \beta_{in} \prod_{j=1}^Q X_j^{h_{ijn}} = 0 \quad i = 1 \dots M \quad (3.30)$$

Donde los índices de los términos positivos y negativos dominantes son p y n .

A partir de el subsistema (3.30) se toman logaritmos y se reorganiza la ecuación:

$$\log(\alpha_{ip}) + \sum_{j=1}^Q g_{ijp} \log(X_j) = \log(\beta_{in}) + \sum_{j=1}^Q h_{ijn} \log(X_j) \quad (3.31)$$

Finalmente, esto resultan en un sistema lineal del que se puede analizar su estabilidad a través de evaluar los eigenvalores de su matriz jacobiana. Se pueden tener tantos

3. METODOLOGÍA

subsistemas como combinaciones de términos positivos con términos negativos. Dada la forma de la ecuación (3.31), los valores g_{ijp} y h_{ijn} , que corresponden a los exponentes en las ecuaciones, deben permanecer fijos para resolver el sistema.

Como esta aproximación implica que uno de los términos adquiere mayor valor que los demás (para los positivos y para los negativos), resulta en desigualdades de las que también puede tomarse el logaritmo y resultar en un sistema lineal. Gracias a esto se pueden obtener los límites de los fenotipos y los parámetros de estas regiones quedan restringidos a aquellos que cumplan con las desigualdades. Cabe mencionar que la aproximación es mejor lejos de estos límites y peor cerca de ellos.

El modelo se implementó en la Versión 3 de *Design Space Toolbox* (DST) (37). El programa está diseñado para que de manera interactiva el usuario pueda introducir su sistema en la forma de GMA y posteriormente producir los análisis deseados. La forma GMA de nuestro modelo es la siguiente:

$$\dot{A} = K_1 \alpha K_B^{n_B} X_1^{-1} + K_1 A^{n_A} X_2^{-1} - K_1 \alpha K_B^{n_B} X_1^{-1} A^{n_A} X_2^{-1} - A \quad (3.32)$$

$$\dot{B} = K_2 \beta + K_2 A^n Y^{-1} - K_2 \beta A^n Y^{-1} - B \quad (3.33)$$

$$X_1 = K_B^{n_B} + B^{n_B} \quad (3.34)$$

$$X_2 = A^{n_A} + K_A^{n_A} K_B^{-n_B} B^{n_B} + K_A^{n_A} \quad (3.35)$$

$$Y = A^n + K^n \quad (3.36)$$

Aquí X_1 , X_2 y Y son variables auxiliares mientras que \dot{A} y \dot{B} son las ecuaciones diferenciales del modelo. Las variables auxiliares contienen a los denominadores para que el sistema pueda escribirse como una suma de productos.

3.1.3.2. Exploración basada en parámetros

El muestro Monte Carlo se hizo a partir de una distribución uniforme. Para cada uno de los parámetros se tomaron valores de la distribución uniforme con rangos según se indiquen en la sección de resultados; para algunos casos el muestreo fue sobre solo ciertos parámetros, dejando los demás fijos. Para el análisis de estabilidad se utilizó el mismo algoritmo descrito en la sección 3.1.2, solo que de manera recursiva para el número de conjuntos de parámetros deseados.

3.2. Simulaciones Estocásticas del sistema

Tomando en cuenta el papel del ruido bioquímico en el sistema se realizaron simulaciones estocásticas utilizando el algoritmo de Gillespie (13). En la Tabla 3.1 se muestran las reacciones del modelo desarrollado en este trabajo, así como sus propensidades.

En donde:

Evento	Descripción	Propensidades
$A \longrightarrow A + 1$	Síntesis de A	P_1
$A \longrightarrow A - 1$	Degradación de A	P_2
$B \longrightarrow B + 1$	Síntesis de B	P_3
$B \longrightarrow B - 1$	Degradación de B	P_4

Tabla 3.1: Reacciones del modelo - Eventos bioquímicos con sus respectivas propensidades

$$P_1 = K_1 \left(\alpha \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} + (1 - \alpha) \frac{K_B^{n_B}}{B^{n_B} + K_B^{n_B}} \right) \frac{A^{n_A}}{A^{n_A} + \frac{K_A^{n_A}}{K_B^{n_B}} (B^{n_B} + K_B^{n_B})} \quad (3.37)$$

$$P_2 = K_2 A \quad (3.38)$$

$$P_3 = K_3 (\beta + (1 - \beta) \frac{A^n}{A^n + K^n}) \quad (3.39)$$

$$P_4 = K_4 B \quad (3.40)$$

El algoritmo se implementó en C++ y los resultados se graficaron en Wolfram Mathematica o en R.

3.3. Modelo evolutivo

El circuito se evolucionó en una población celular sometida a un ambiente fluctuante; para ello se utilizó el algoritmo desarrollado en Gómez-Schiavon et al. (16) con las modificaciones pertinentes.

3.3.1. Antecedentes

Brevemente, el algoritmo original evolucionaba una población con un circuito de retroalimentación positiva de un solo gen. El ambiente en el que se encontraba la población fluctuaba entre dos estados, cada uno de ellos asociados a un número de proteína óptimo. Cada miembro de la población celular contaba con ciertos parámetros (i.e. genotipo) y con ellos se simulaban sus dinámicas de expresión (i.e. fenotipo). Si el número de proteínas expresadas por la célula era cercano al número de proteínas óptimo para el ambiente en el que se encontraba, era más probable que la célula se seleccionara para la siguiente generación. El modelo permitía mutaciones, de manera que los miembros de la población podían adaptarse al ambiente fluctuante a través de mutaciones (i.e. estrategia de adaptación genética) o por medio de *state switching* (i.e. estrategia en la que se puede cambiar entre estados fenotípicos manteniendo un mismo genotipo). En la siguiente sección se describe el modelo evolutivo utilizado en este trabajo.

3.3.2. Descripción

La población es haploide y asexual y simula un modelo Wright-Fisher (i.e. generaciones no superpuestas y tamaño de la población fijo) con modificaciones que se explican a continuación. Cada miembro de la población tiene el circuito génico de retroalimentación positiva y negativa desarrollado en este trabajo; el genotipo de cada individuo está dado por los parámetros biofísicos del modelo, es decir por los valores de $\{K_1, \alpha, K_A, n_A, K_B, n_B, K_2, K_3, \beta, K, n, K_4\}$. Diferentes conjuntos de parámetros pueden producir dinámicas particulares en las proteínas; de tal forma que al cambiar las constantes de disociación, las tasas de síntesis máxima, síntesis basal y cooperatividad, las dinámicas del circuito también pueden cambiar y por consiguiente su fenotipo.

Por simplicidad, el fenotipo que estará bajo selección durante toda la simulación será solo el número de una de las proteínas. El ambiente en el que se encuentra la población fluctúa cada f generaciones entre $E = \{H, L\}$ y a cada ambiente le corresponde un número óptimo de proteína: $P^{(H)} = 80$ y $P^{(L)} = 20$, respectivamente y en donde P puede ser A o B. Llamo época al número de generaciones en que la población permanece en el mismo ambiente, mientras que un ciclo ambiental consiste de dos épocas en ambientes distintos, como se indica en la Figura 3.3 (en donde se ejemplifica la dinámica cuando la proteína A es la que se encuentra bajo selección). El modelo evolutivo se simula durante 10,000 generaciones; la primera generación consiste en una población isogénica (i.e. con el mismo genotipo). En cada generación se realizan los siguientes pasos, esquematizados en la Figura 3.4, ejemplificados para el caso en el que la proteína A es la que se encuentra bajo selección.

1. Simulación de las dinámicas de expresión génica

Para cada miembro de la población y de acuerdo a su genotipo, se simulan estocásticamente las dinámicas de expresión génica del circuito durante el ciclo de

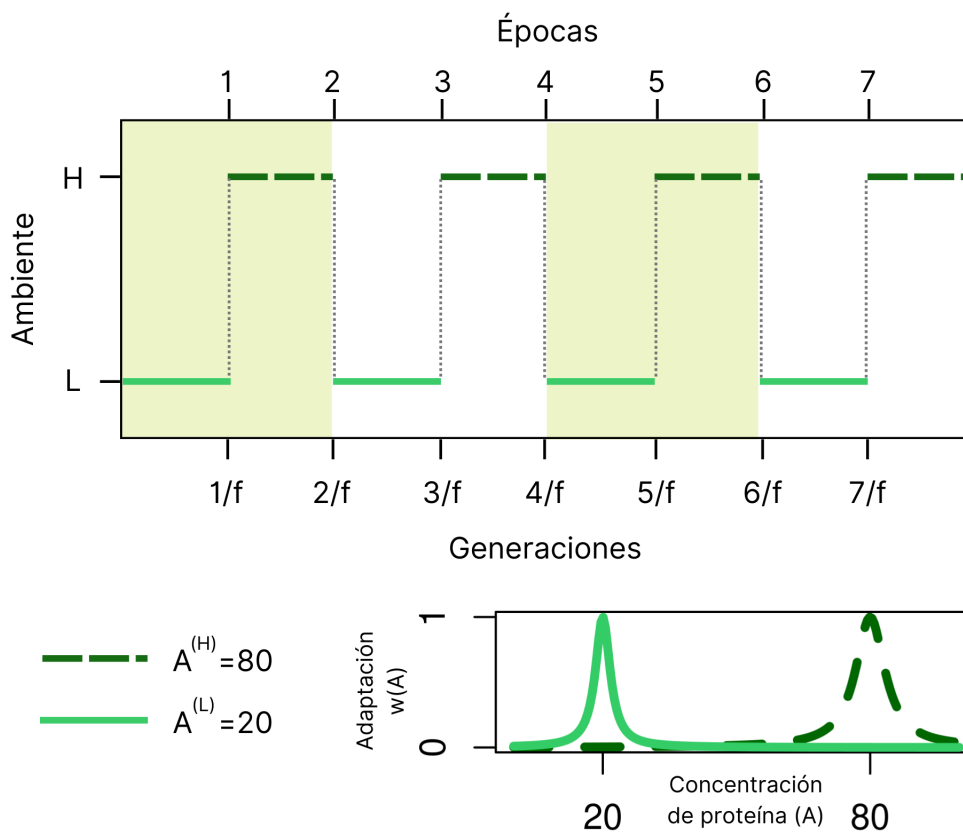


Figura 3.3: Esquema de la dinámica ambiental con selección sobre la proteína A - El ambiente fluctúa cada $\frac{1}{f}$ generaciones entre $E = \{H, L\}$. El número de generaciones que permanecen en el mismo ambiente corresponden a una época. Dos épocas en ambientes distintos ocurren cada $\frac{2}{f}$ y corresponden a un ciclo ambiental. En cada ambiente se selecciona para un número de proteínas distinto, correspondientes a $A^{(H)} = 80$ y $A^{(L)} = 20$, cuando se está en el ambiente alto (H) y bajo (L), respectivamente.

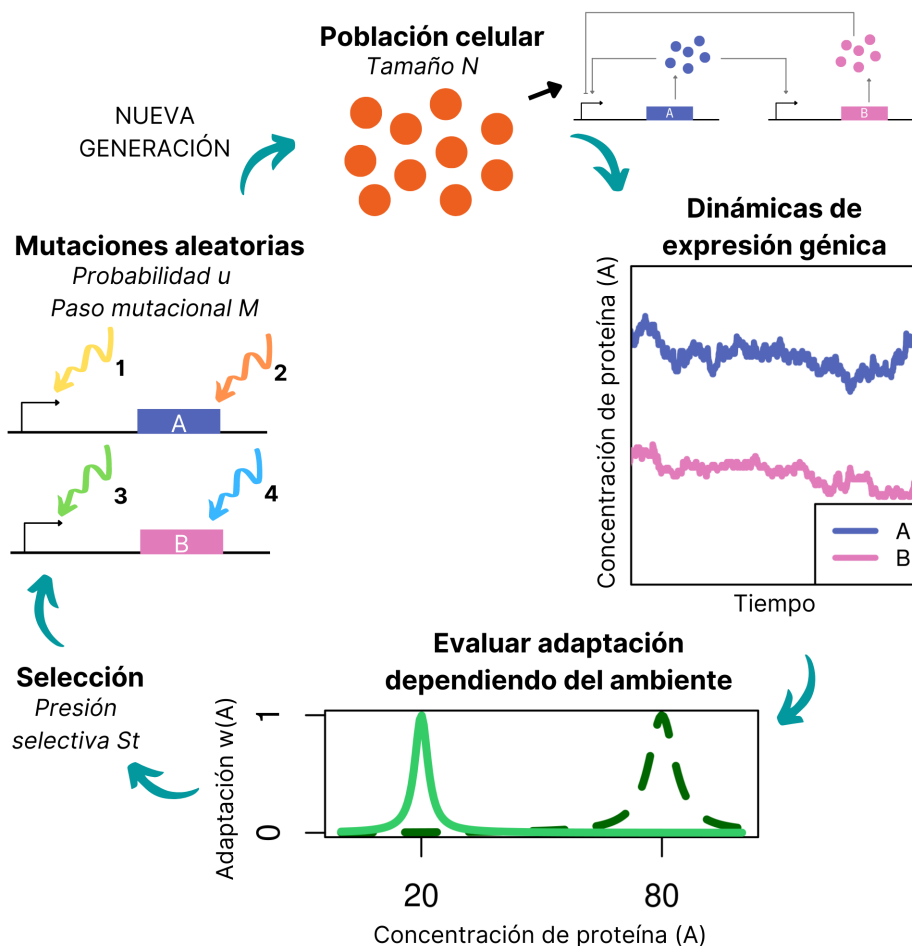


Figura 3.4: Esquema del ciclo generacional con selección sobre la proteína A - Cada integrante de la población de tamaño N tiene el circuito de regulación génica y un genotipo propio. En cada individuo se simulan las dinámicas de expresión estocásticas del circuito. Se evalúa el valor de la proteína A al final de la simulación y se le asigna un valor de adaptación $w(A)$ a la célula, en el que el valor óptimo de A depende del ambiente. La siguiente generación se obtiene a través de selección de torneo (donde la presión selectiva está relacionada con el número de miembros del torneo S_t) de acuerdo al valor de $w(A)$. Las células de la siguiente generación tienen una probabilidad u de mutar en cuatro posibles escenarios y con un paso mutacional M .

vida de la célula. En el modelo se define este ciclo como 4 unidades de tiempo y el número inicial de las proteínas A y B son heredados de la célula antecesora.

2. Evaluar adaptación de acuerdo al ambiente

Al final del ciclo de vida de la célula, el valor de expresión de la proteína P es evaluado en una función de adaptación $w(P)$. La función de adaptación es una variación de la distribución de Lorentz y está dada por la ecuación (3.41). En donde $\Gamma = 0.2 \cdot P^{(E)}$, siendo $P^{(E)}$ el valor óptimo de proteína P ($P^{(H)} = 80$ o $P^{(L)} = 20$) dependiendo del ambiente en el que se encuentre la población ($E = \{H, L\}$); y siendo P el valor expresado por la célula. De esta forma, mientras P sea más cercano al óptimo $P^{(E)}$, el valor de adaptación de la célula será más cercano a 1.

$$w(P) = \frac{\Gamma}{\Gamma + (P - P^{(E)})^2} \quad (3.41)$$

La función de adaptación está ilustrada en la Figura 3.3.

3. Selección de acuerdo a adaptación

Para seleccionar a la siguiente generación se utilizó el algoritmo de selección por torneo. De la población se escogen S_t número de células que conforman el torneo, de ellas la célula con el valor de adaptación mayor es clonada para la siguiente generación. Se hacen N número de torneos con reemplazo para conservar el tamaño poblacional. El parámetro S_t determina la presión selectiva: $S_t = 1$ indican una presión selectiva nula donde la deriva génica juega el rol más importante ya que los individuos que pasen a la siguiente generación serán seleccionados al azar; $S_t = N$ indica que solo el individuo mejor adaptado será clonado. De manera que valores pequeños de S_t indican una presión selectiva baja mientras que valores altos de S_t una presión selectiva alta.

4. Permitir mutaciones aleatorias en las células seleccionadas

Se permiten mutaciones sobre todos los parámetros salvo por las constantes de degradación, que por simplicidad permanecen fijas en los valores $K_2 = 1$ y $K_4 = 1$. En cada célula existe una probabilidad u de que una mutación ocurra en el circuito. Si se lleva acabo una mutación existen cuatro regiones en las que puede ocurrir, todas con la misma probabilidad: el promotor de A (1), la región codificante de A (2), el promotor de B (3) y la región codificante de B (4). Cuando ocurre una mutación en el promotor de una proteína puede verse afectada la expresión génica, la afinidad de los factores de transcripción con dicho promotor y la manera en la que los factores de transcripción interactúan entre ellos o con el resto de la célula para poder regular al promotor. Por otra parte, mutaciones en la región codificante pueden afectar directamente a la proteína y cómo interactúa con el promotor, con el resto de la célula (incluidas con otras moléculas de la misma proteína). Por ejemplo, pudiera darse el caso en el que originalmente un factor de transcripción necesitara ser fosforilado en tres sitios para poder unirse a su promotor y que una mutación en el promotor facilite que con solo dos sitios fosforilados el factor de transcripción cumpla su función.

3. METODOLOGÍA

Considerando lo anterior, al ocurrir mutaciones en el promotor de A y B los parámetros afectados serán las tasas de síntesis máxima (K_1 para el promotor de A; K_3 para el promotor de B), la actividad basal relativa (α para el promotor de A; β para el promotor de B), las constantes de disociación de las proteínas por los promotores (K_A y K_B para el promotor de A; K para el promotor de B) y la cooperatividad del proceso de transcripción (n_A y n_B para el promotor de A; n para el promotor de B). Al ocurrir mutaciones en las regiones codificantes los parámetros afectados serán K_A , K_B , n_A y n_B para la proteína A; K y n para la proteína B. Por simplicidad, no se consideran mutaciones que causen pérdida de función y solo una mutación puede ocurrir por célula. Los parámetros que se ven afectados de acuerdo a la región que sufre la mutación están resumidos en la Tabla 3.2.

Escenario	Región	Parámetros afectados
1	Promotor de A	$K_1, \alpha, K_A, n_A, K_B, n_B$
2	Región codificante de A	K_A, n_A, K, n
3	Promotor de B	K_3, β, K, n
4	Región codificante de B	K_B, n_B

Tabla 3.2: Escenarios mutacionales - Parámetros afectados de acuerdo a la región en la que ocurre la mutación.

Si la mutación ocurre, los parámetros se actualizan de la siguiente manera:

$$X'_i \leftarrow X_i M^r \quad (3.42)$$

Donde X_i corresponde al valor previo del parámetro mutado y el número de parámetros que mutarán depende del escenario mutacional seleccionado. M corresponde al paso mutacional, es decir el máximo cambio que puede sufrir el parámetro, y $r \sim U(-1, 1)$ sigue una distribución uniforme entre -1 y 1. Este esquema mutacional permite co-variación en los parámetros afectados de acuerdo al escenario seleccionado, de manera que en escala logarítmica las posibles mutaciones se encuentran uniformemente distribuidas en un cubo de n -dimensiones en un rango definido por M .

Los parámetros pueden adquirir valores restringidos de acuerdo a los que consideramos biológicamente factibles, y se basan en los límites utilizados en Gómez-Schiavon et al. (16). Si las mutaciones llevan al parámetro a un valor superior o inferior a los establecidos por los límites, el parámetro adquiere el valor máximo o mínimo posible, según corresponda. Los rangos de valores que pueden adquirir

son: para $K_1 \in [0.01, 1000]$, para $K_3 \in [0.01, 1000]$ o $\in [0.01, 150]$ según se indica en la sección 3.3.3, para α y $\beta \in [0.001, 0.999]$, para K_A, K_B y $K \in [0.01, 100]$ y para n, n_A y $n_B \in [0.01, 20]$.

5. Pasar a la siguiente generación y repetir el ciclo hasta que se llegue a la última generación. Si la generación es la última de la época, actualizar el ambiente.

El algoritmo original se implementó en C++ por Gómez-Schiavon et al (16). El código modificado descrito aquí se encuentra en el Apéndice 2 (A.2); requiere de la librería que se encuentra en el Apéndice 3 (A.3) que define el circuito de regulación y varias funciones utilizadas en el código principal.

3.3.3. Diseño Experimental de las Simulaciones Evolutivas

Todas las poblaciones se simularon por 10000 generaciones, iniciando siempre en el ambiente bajo en donde el óptimo es $P \sim 20$. Se realizaron dos tipos de experimentos con el algoritmo descrito en la sección anterior: el primero seleccionando sobre la concentración de la proteína A y el segundo sobre la concentración de la proteína B. Considerar ambas proteínas permite explorar si el circuito puede tener dinámicas evolutivas distintas de acuerdo a la proteína que esté bajo selección. Para las simulaciones que seleccionan sobre la proteína A se utilizó el rango de valores para K_3 de $[0.01, 150]$; para la selección sobre la proteína B el rango de K_3 utilizado fue $[0.01, 1000]$. Esto fue porque en las simulaciones en las que A se encontraba bajo presión selectiva, al conservar el rango original de K_3 se observaba que este parámetro continuaba subiendo y la población no parecía adaptarse.

Además, para investigar si el genotipo inicial podía influir en dichas dinámicas, para cada uno de los experimentos se realizaron las simulaciones desde 4 genotipos iniciales distintos: (1) estado estacionario alto en $P \sim 80$, (2) estado estacionario bajo en $P \sim 20$, (3) oscilaciones con picos y valles en $P \sim 80$ y $P \sim 20$, respectivamente y (4) biestabilidad con estados estacionarios en $P \sim 80$ y $P \sim 20$. En donde P se refiere a A o B según el experimento: para las simulaciones que seleccionan sobre A, P corresponde a A, mientras que para las simulaciones que seleccionan sobre B, P corresponde a B. Simulaciones estocásticas de los genotipos iniciales para cada experimento se encuentran la Figura 3.5.

Para cada experimento se simularon diferentes frecuencias del cambio ambiental (f) y pasos mutacionales (M). Nos enfocamos inicialmente en ellos porque en Gómez-Schiavon et al. (15) se observó que cambios en M y f llevaban a un cambio gradual en la proporción de las diferentes estrategias adaptativas. El resto de los parámetros se escogió basándose en Gómez-Schiavon et al. (15), eligiendo parámetros en los que las estrategias estudiadas en dicho trabajo coexistían en las poblaciones, bajo el supuesto de que dicha elección de parámetros aquí permitiría la competencia entre distintas estrategias,

3. METODOLOGÍA

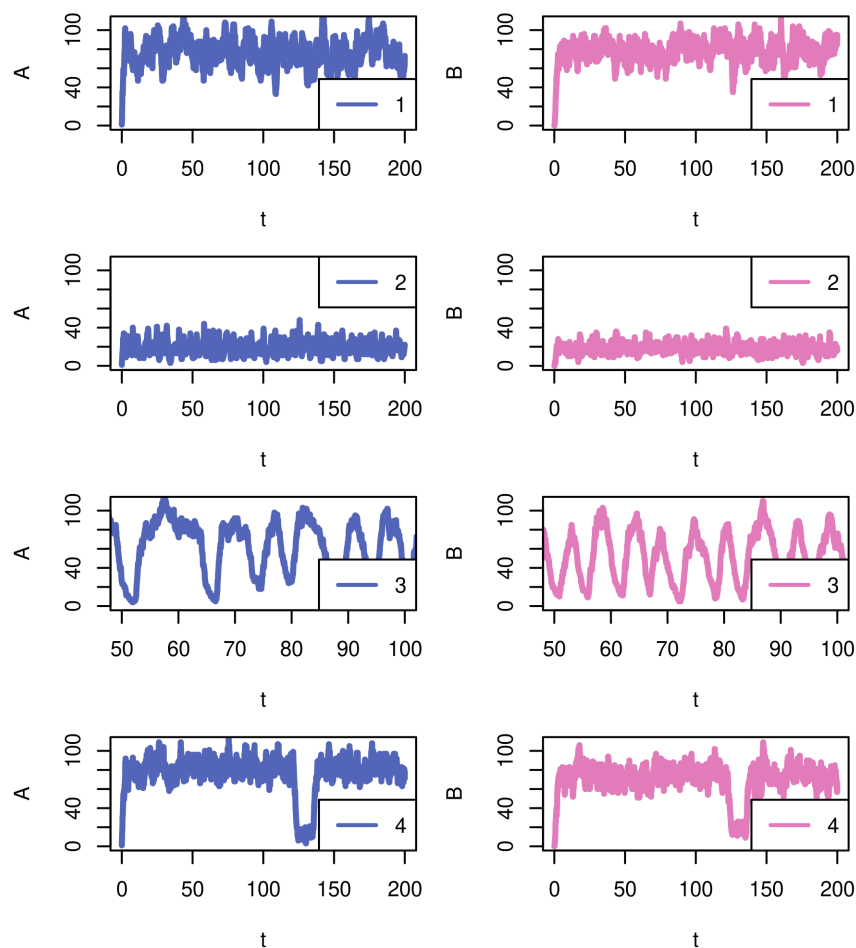


Figura 3.5: Genotipos iniciales - Parámetros para los genotipos iniciales en la selección sobre A (izquierda): **(1)** $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 50$, $n = 4$, $K_4 = 1$. **(2)** $K_1 = 100$, $\alpha = 0.25$, $K_A = 30$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 40$, $n = 4$, $K_4 = 1$ **(3)** $K_1 = 100$, $\alpha = 0.25$, $K_A = 6$, $K_B = 10$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 120$, $\beta = 0.01$, $K = 40$, $n = 4$, $K_4 = 1$. **(4)** $K_1 = 84$, $\alpha = 0.6$, $K_A = 25$, $K_B = 15$, $n_A = 16$, $n_B = 7$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 24$, $n = 2$, $K_4 = 1$. Parámetros para los genotipos iniciales en la selección sobre B (derecha): **(1)** $K_1 = 50$, $\alpha = 0.5$, $K_A = 3$, $K_B = 10$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 20$, $n = 5$, $K_4 = 1$. **(2)** $K_1 = 50$, $\alpha = 0.5$, $K_A = 30$, $K_B = 20$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 50$, $\beta = 0.01$, $K = 20$, $n = 5$, $K_4 = 1$ **(3)** $K_1 = 80$, $\alpha = 0.25$, $K_A = 10$, $K_B = 15$, $n_A = 5$, $n_B = 5$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.01$, $K = 20$, $n = 3$, $K_4 = 1$. **(4)** $K_1 = 84$, $\alpha = 0.6$, $K_A = 25$, $K_B = 15$, $n_A = 16$, $n_B = 7$, $K_2 = 1$, $K_3 = 80$, $\beta = 0.01$, $K = 24$, $n = 2$, $K_4 = 1$.

concentrándonos por el momento en estudiar el efecto de la variación de M y f . Los parámetros utilizados se encuentran en la Tabla 3.3.

Parámetro evolutivo	Valor(es)
Tamaño poblacional (N)	4000
Número de células que entran en la selección de torneo (S_t)	15
Pobabilidad de que ocurra una mutación (u)	0.03
Frecuencia del cambio ambiental (f)	{0.5, 0.2, 0.1, 0.02}
Paso mutacional (M)	{1.1, 2, 3}

Tabla 3.3: Parámetros utilizados para las simulaciones evolutivas - Para todas las simulaciones, los parámetros N , S_t y u tienen el mismo valor. Una simulación se hizo para cada combinación de los valores de f y M .

3.3.4. Análisis de las simulaciones evolutivas

Para estudiar cómo la población lidia con el ambiente fluctuante se utilizaron principalmente dos estrategias. La primera consiste en analizar los valores de adaptación promedio de la población en cada generación junto con los valores promedio de las concentraciones de las proteínas del circuito, igualmente en cada generación; permitiendo saber si después de las 10000 generaciones simuladas la población logra adaptarse.

La segunda consiste en estudiar el comportamiento (i.e. la dinámica del circuito) de los linajes celulares que sobreviven durante al menos un ciclo ambiental completo. Esto nos permite detectar estrategias capaces de lidiar con las fluctuaciones, con mayores valores de adaptación, y así caracterizar las estrategias adaptativas exitosas. En el resto del trabajo nos referimos a los linajes como a la línea de descendencia de una célula ancestral que permanece en la población durante al menos un ciclo ambiental.

Para asegurarnos que las estrategias son capaces de lidiar con las fluctuaciones ambientales, consideramos solo los linajes que sobreviven dos ciclos ambientales completos, usando los genotipos del primer ciclo para definir la estrategia adaptativa utilizada (Figura 3.6). Esto nos permite asegurar que cada uno de los genotipos observados en el primer ciclo tuvo descendientes que han sobrevivido a los dos tipos de transición ambiental (i.e. L a H y H a L). Si por ejemplo nos fijáramos solo en las últimas generaciones del segundo ciclo, los genotipos analizados podrían ser una nueva mutante que no necesariamente podría sobrevivir al siguiente ciclo. Los linajes se tomaron de los últimos 100 ciclos ambientales, para permitir a la población adaptarse al ambiente

3. METODOLOGÍA

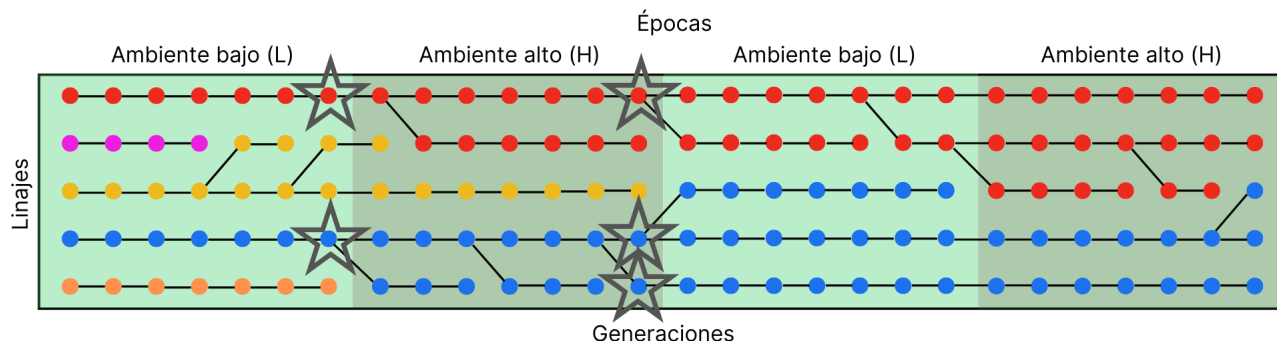


Figura 3.6: Diagrama de la elección de individuos que se analizan para caracterizar las estrategias adaptativas - Las células marcadas con una estrella son aquellas cuyas dinámicas se caracteriza, pues son las de la última generación de cada época del primer ciclo de aquellos linajes que sobreviven dos ciclos ambientales.

fluctuante en primer lugar. En cada ciclo considerado se recabó el número de mutaciones del linaje durante el ciclo ambiental, la dinámica que utilizó el linaje en su última generación en cada uno de los ambientes (determinado con el código descrito en la sección 3.1.2) y el valor de adaptación promedio del linaje en el ciclo ambiental. En la Figura 3.6 se esquematizan las células del linaje que son analizadas.

Las figuras del Capítulo 6 se hicieron en R utilizando la librería *ggplot2*.

Dinámicas del circuito de regulación génica con retroalimentación positiva y negativa

4.1. Solución oscilatoria

Como se muestra en la Figura 4.1, el sistema es capaz de oscilar; la dinámica es la siguiente: al inicio, tanto la proteína A como B están siendo sintetizadas principalmente de manera basal. Dada la ultrasensibilidad de A , al llegar a cierta concentración su síntesis se dispara a valores muy altos, provocando también que la síntesis de B se dispare. Al aumentar la concentración del represor B , el nivel de expresión de A disminuye, causando que la concentración de B también baje. Al disminuir B , A vuelve a aumentar su síntesis y el ciclo se repite.

De acuerdo al análisis de estabilidad, el estado estacionario del conjunto de parámetros elegido para este ejemplo es una espiral inestable. Independientemente de las condiciones iniciales, las trayectorias se verán atraídas a un ciclo límite y por lo tanto el sistema oscilará. En la Figura 4.1 se encuentra el plano fase del sistema, así como sus *nullclines*. El estado estacionario del sistema se encuentra en la intersección de ambas. Aquí, A corresponde al eje x y B al eje y , de manera que A determina el flujo horizontal y B el flujo vertical de las trayectorias. Las trayectorias del sistema en el plano fase, representadas en forma de flechas, muestran que condiciones iniciales lejos del ciclo límite se ven atraídas a él, mientras que las trayectorias más cercanas al estado estacionario se ven repelidas por este y terminan en el ciclo.

4. DINÁMICAS DEL CIRCUITO DE REGULACIÓN GÉNICA CON RETROALIMENTACIÓN POSITIVA Y NEGATIVA

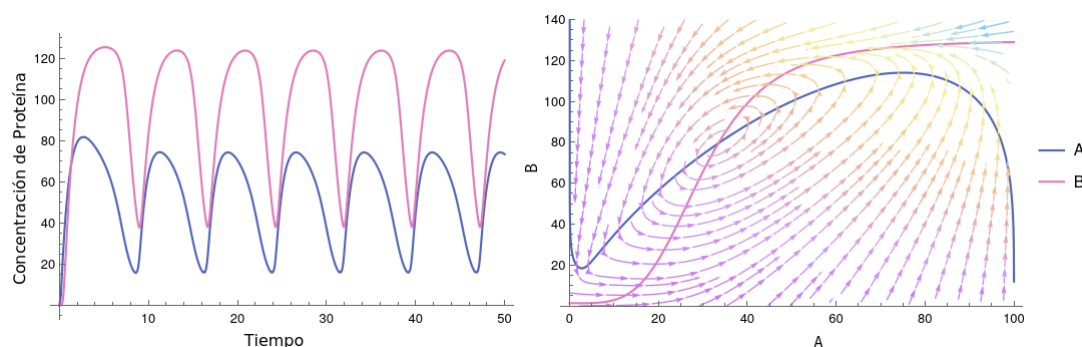


Figura 4.1: Dinámica oscilatoria - Simulación determinista del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 30$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (izquierda). Plano fase con las *nullclines* del sistema (derecha).

4.2. Solución biestable

En la Figura 4.2 se muestra una solución biestable del sistema. En el plano fase se pueden ver los tres estados estacionarios: dos estables y uno inestable. Dependiendo de cuáles sean las condiciones iniciales, las trayectorias se ven atraídas a uno u otro de los puntos atractores. Si la condición inicial se encuentra cerca del punto estable bajo (i.e. en el que ambas proteínas tienen valores menores), la solución se verá atraída a este punto atractor, mientras que si empiezan cerca del estado estacionario alto (i.e. en el que ambas proteínas tienen valores mayores), la solución se dirigirá a este último.

Cabe resaltar que si bien ambos puntos son atractores, las dinámicas presentadas por las proteínas para alcanzarlos no es la misma. Para el punto estable bajo puede verse que la concentración de ambas proteínas se ‘dispara’ y después decrece hasta llegar al estado estacionario; por otra parte, para el punto estable alto no hay ‘disparo’ y ambas proteínas suben suavemente. Estas diferencias también son notables en el mapa fase, las trayectorias llegan de manera más directa al estado estacionario alto y al bajo llegan en forma de espiral. Del análisis de estabilidad sobre los puntos de equilibrio resulta que el punto estable bajo es una espiral estable mientras que el alto un nodo estable.

4.3. Caracterización del espacio de parámetros

Conocer las regiones del espacio de parámetros es relevante para tener en cuenta las dinámicas posibles del sistema así como su robustez; es decir si existen más comportamientos además de los descritos en las secciones anteriores y si el modelo es capaz de conservar sus dinámicas ante cambios en el espacio de parámetros. Para hacer esta

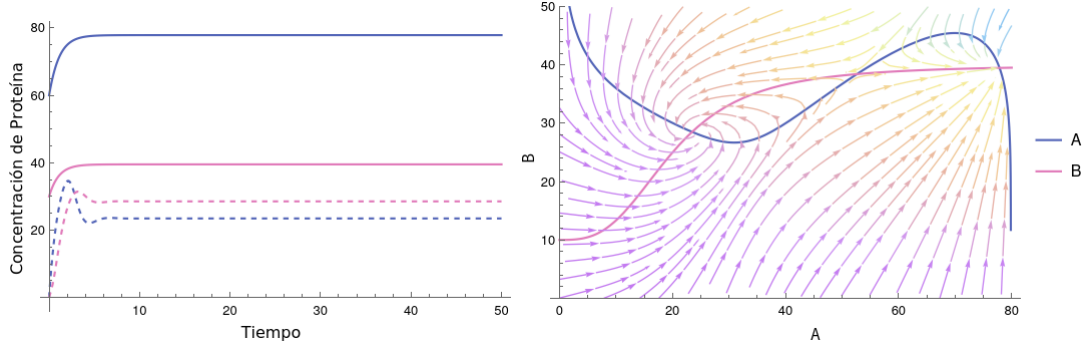


Figura 4.2: Dinámica biestable - Simulación determinista del circuito con parámetros: $K_1 = 80$, $\alpha = 0.5$, $K_A = 40$, $K_B = 30$, $n_A = 8$, $n_B = 6$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.25$, $K = 20$, $n = 3$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (líneas punteadas) y $A = 60$ y $B = 30$ (líneas sólidas) (izquierda). Plano fase con las *nullclines* del sistema (derecha).

caracterización se utilizó el software de *Design Space Toolbox* (DST)(37). Este análisis resultó en un listado de 18 fenotipos resumidos en la Tabla 4.1.

Caso	Eigenvalores	Número de fenotipos
1	2 positivos con componente compleja	1
2	0 positivos con componente compleja	1
3	0 positivos sin componente compleja	11
4	1 positivo sin componente compleja	5

Tabla 4.1: Fenotipos del sistema - Cada caso está definido por el tipo de estabilidad para su estado estacionario; la matriz jacobiana del sistema tiene dos eigenvalores, aquí se reporta el número de ellos que son positivos y si tienen componente compleja. Para cada caso se indica el número de fenotipos. Datos del modelo con $n_A = 4$, $n_B = 4$ y $n = 4$.

Los casos 1 y 2, al tener componente compleja, corresponden a fenotipos con espirales. En el caso 1 la espiral es inestable, pues tiene eigenvalores positivos, mientras que para el caso 2 es estable, ya que sus eigenvalores son negativos. El caso 3 corresponde a un nodo atractor y el caso 4 a un *saddle point*. Este último tipo de estabilidad surge cuando los eigenvalores tienen signos contrarios, resultando en que el estado estacionario sea inestable. Como cada fenotipo representa la solución sobre un solo estado estacionario,

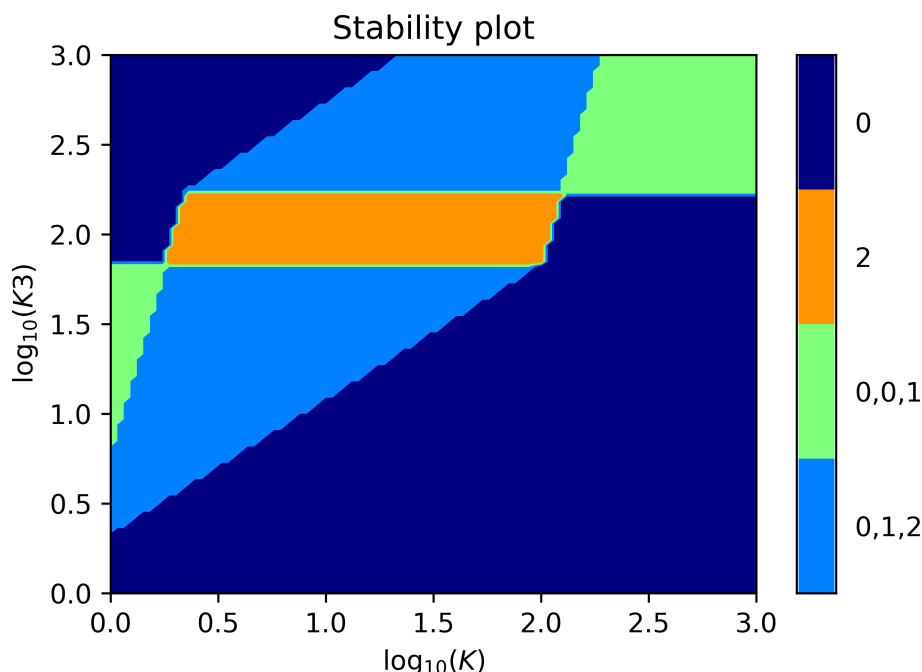


Figura 4.3: Gráfica de estabilidad de acuerdo a *DST* - Estabilidad de las regiones de parámetros de acuerdo a cambios en K y K_3 . Las clasificaciones de estabilidad están descritas en el texto principal. El conjunto de parámetros referencia para esta figura: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; este conjunto se encuentra en la región 2 de la figura.

comportamientos con más de uno como la biestabilidad se encuentran en la intersección de tres fenotipos: dos con puntos de equilibrio estables y uno inestable (e.g. dos fenotipos del caso 3 y uno del caso 4). Con el propósito de explorar la región de parámetros alrededor de la región oscilatoria se tomó un conjunto de parámetros para los que el sistema oscila y con *DST* se hizo una gráfica que muestra cómo cambia la estabilidad del sistema ante cambios en los parámetros K y K_3 (Figura 4.3). Según esta estimación las regiones tienen la siguiente estabilidad: (0) punto de equilibrio atractor, (2) espiral inestable (donde se encuentra el conjunto de parámetros de referencia), (0,0,1) dos puntos de equilibrio atractores y un repulsor, (0,1,2) punto de equilibrio atractor, un repulsor y una espiral inestable. Las aproximaciones de *DST* están basadas en el análisis de estabilidad de un conjunto de parámetros representativo de la región por lo que validarlas se hizo una exploración basada en parámetros (sección 3.1.3.2), dejando todos los parámetros fijos salvo por K_3 y K . Los resultados se muestran en la Figura 4.4.

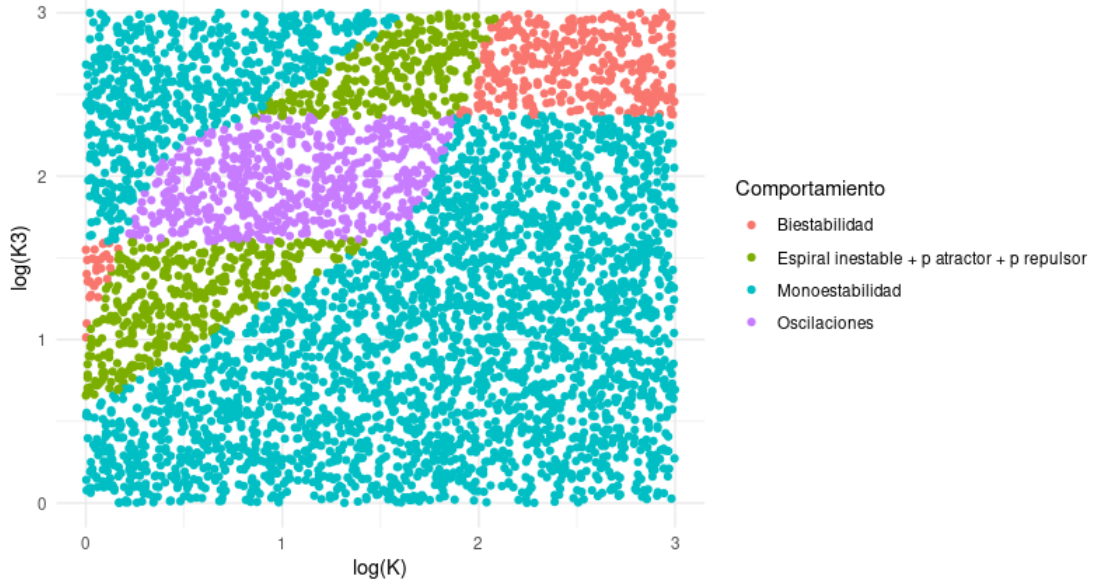


Figura 4.4: Gráfica de estabilidad de acuerdo a la exploración basada en parámetros - Muestreo de 5,000 conjuntos de parámetros. El rango de valores de K y K_3 es $[0,3]$ en escala logarítmica. Parámetros fijos: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $\beta = 0.01$, $n = 4$, $K_4 = 1$.

Al comparar la Figura 4.3 de *DST* con la Figura 4.4, se puede ver que en general el patrón se mantiene pero algunos conjuntos de parámetros para los que *DST* predice cierta estabilidad, realmente presentan un comportamiento distinto. Por ejemplo, en la Figura 4.3 pareciera ser que con los parámetros $\log_{10}(K_3) = 0.5$ y $\log_{10}(K) = 0$, el sistema se encontraría en la región $(0,1,2)$, pero al ver parámetros cercanos a estos valores en la Figura 4.4, estos presentan un comportamiento monoestable.

Salvo por la región $(0,1,2)$ el comportamiento esperado es conocido (e.g. se espera que en la región fenotípica $(0,0,1)$ el sistema sea biestable), por ello se analizó un caso en dicha región. En la Figura 4.5 se muestra la solución numérica y el mapa fase de un conjunto de parámetros en la región fenotípica $(0,1,2)$. En el mapa fase se pueden observar los tres puntos fijos con la estabilidad predicha; a la izquierda está la espiral inestable, a la derecha el nodo estable y entre ellos el nodo inestable. En la simulación determinista se observa como el sistema llega a su estado estacionario, además de un disparo inicial en la concentración de la proteína A.

4. DINÁMICAS DEL CIRCUITO DE REGULACIÓN GÉNICA CON RETROALIMENTACIÓN POSITIVA Y NEGATIVA

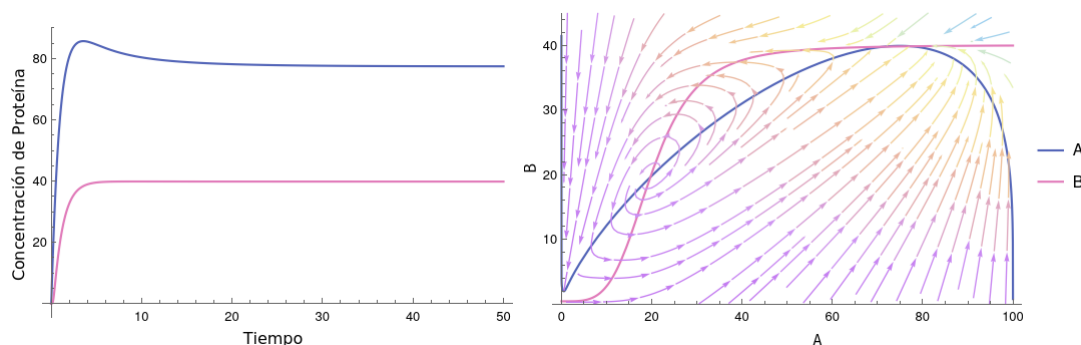


Figura 4.5: Dinámica en la región (0,1,2) - Simulación determinista del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (izquierda). Plano fase con las *nullclines* del sistema (derecha).

4.4. Discusión

Como se esperaba, el sistema desarrollado es capaz de presentar dinámicas oscilatorias y biestables. Las oscilaciones que genera este circuito corresponden a la forma de oscilaciones de relajación descritas en el Capítulo 2. Las oscilaciones de la Figura 4.1 tienden a pasar más tiempo en concentraciones altas para después bajar y volver a subir precipitadamente. Con otros parámetros también se han observado oscilaciones en las que el patrón es el opuesto, pero que también corresponde a oscilaciones de relajación, en donde tienden a pasar más tiempo en concentraciones bajas y dispararse y bajar precipitadamente, formando picos. El análisis de estabilidad del sistema con parámetros que generan biestabilidad en la Figura 4.2 permitió observar que las dinámicas de las trayectorias para llegar a los dos estados estacionarios pueden ser distintas y aunque a lo largo del tiempo terminan manteniendo una concentración constante, las que presentan la espiral estable generan antes un pequeño disparo en la concentración.

El análisis del espacio de parámetros con *DST*, además de lograr confirmar que las dinámicas de interés son robustas ante cambios en sus parámetros, permitió encontrar conjuntos de parámetros que presentan comportamientos que no se habían considerado inicialmente (i.e. región (0,1,2)). Se comprobó que esta herramienta puede predecir adecuadamente los comportamientos del sistema, pero que como los desarrolladores indican, va perdiendo precisión cuando se encuentra en los límites de los fenotipos.

Efecto del Ruido Bioquímico en las dinámicas del circuito

5.1. Las oscilaciones son robustas ante el ruido bioquímico

La dinámica oscilatoria del modelo es robusta ante el ruido bioquímico. Como se puede ver en la Figura 5.1, los parámetros utilizados para la Figura 4.1 presentan oscilaciones aún ante las perturbaciones del ruido. Sus amplitudes y periodo conservan valores similares aunque las amplitudes de ambas proteínas son mayores en la simulación estocástica y el periodo menos preciso.

5.2. Oscilaciones inducidas por ruido bioquímico

El ruido bioquímico también produce oscilaciones cuando de manera determinista el sistema llega a un estado estacionario. Para los parámetros utilizados en la Figura 5.2, de acuerdo al análisis de estabilidad, el punto de equilibrio es una espiral atractora. Por ello, al integrar de manera numérica las ecuaciones diferenciales hay fluctuaciones en la expresión de las proteínas antes de llegar a su estado estable. Por otra parte, en la simulación estocástica estas fluctuaciones son ampliadas y el resultado es una dinámica oscilatoria.

Como se vio en el Capítulo 2, en la literatura ya se han reportado casos de oscilaciones inducidas por ruido cuando el análisis de estabilidad reporta un estado estacionario estable; dichas oscilaciones tienden a presentar amplitudes similares a las que se observan en la región oscilatoria (determinista) cercana al punto de bifurcación. Para explorar si esto también se observaba en nuestro modelo se procedió a cambiar solo el parámetro K del valor $K = 25$ utilizado en la simulación de la Figura 5.2 a $K = 28$; el sistema con $K = 28$ presenta un ciclo límite (su análisis de estabilidad resultó en una espiral inesta-

5. EFECTO DEL RUIDO BIOQUÍMICO EN LAS DINÁMICAS DEL CIRCUITO

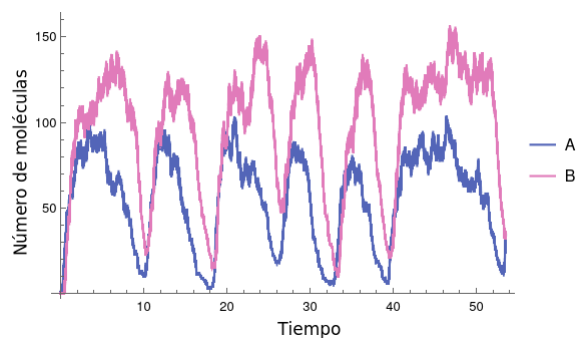


Figura 5.1: Dinámica oscilatoria con ruido bioquímico - Simulación estocástica del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 5$, $K_B = 10$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 130$, $\beta = 0.01$, $K = 30$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$.

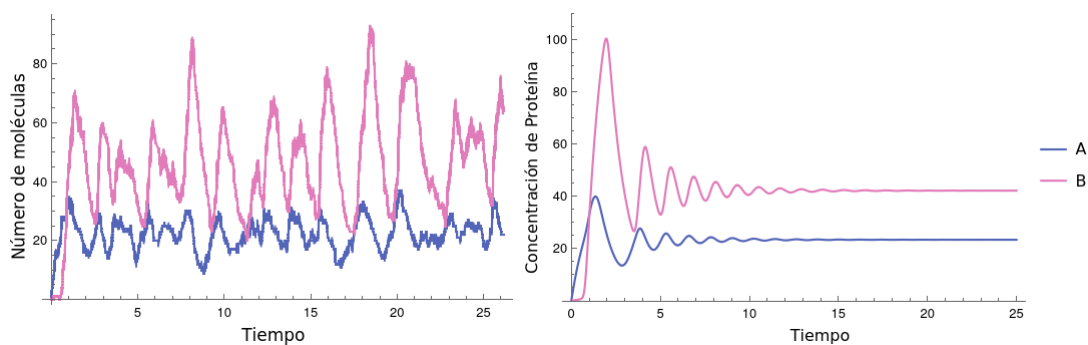


Figura 5.2: Oscilaciones inducidas por ruido bioquímico - Simulaciones estocástica (izquierda) y determinista (derecha) del circuito con parámetros: $K_1 = 90$, $\alpha = 0.5$, $K_A = 30$, $K_B = 40$, $n_A = 8$, $n_B = 5$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.01$, $K = 25$, $n = 15$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$.

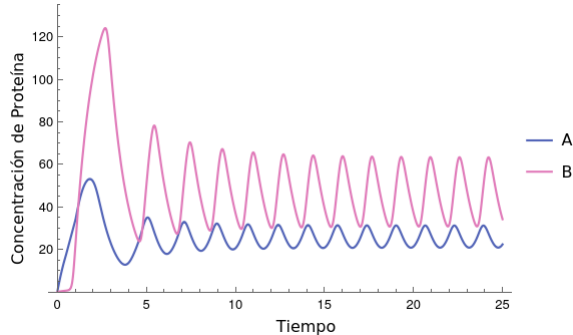


Figura 5.3: Oscilaciones cerca de un punto de bifurcación - Simulación determinista del circuito con parámetros: $K_1 = 90$, $\alpha = 0.5$, $K_A = 30$, $K_B = 40$, $n_A = 8$, $n_B = 5$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.01$, $K = 28$, $n = 15$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$.

ble), de manera que hay oscilaciones en la simulación determinista (Figura 5.3). Tanto en las oscilaciones estocásticas de la Figura 5.2 como en las deterministas de la Figura 5.3, la proteína B alcanza valores alrededor de 60 en sus picos y la proteína A alrededor de 30; además de que el periodo es similar. Las oscilaciones de la simulación estocástica analizada aquí tiende a conservar propiedades de las oscilaciones deterministas que se encuentran cerca en el espacio de parámetros.

5.3. *Switching* biestable

En acuerdo con el Capítulo 2, un sistema biestable con ruido bioquímico es un *switch* biestable en el que el sistema puede cambiar entre los dos estados estables. El resultado de simular de manera estocástica el modelo con los mismos parámetros que los utilizados en la Figura 4.2 se encuentra en la Figura 5.4. Se puede ver que la proteína A empieza en valores cercanos a 30, después de algunas fluctuaciones causadas por el ruido se ‘dispara’ su síntesis y tiende a estar ahora en valores cercanos a 80. A lo largo de la simulación el número de proteínas va cambiando entre estos dos estados. En la Figura 5.4 se muestran los histogramas de la simulación para ambas proteínas. Para el factor de transcripción A , es claro que la distribución es bimodal y que cada moda corresponde a cada uno de los estados estacionarios estables (véase en la Figura 4.2). Por otra parte, tanto en la simulación estocástica como en el histograma, las características del *switching* biestable no parecen ser tan claras en la proteína B : las fluctuaciones no aparentan un *switch* y la distribución parece tener una sola moda. Si se consideran los estados estacionarios estables de B , estos tienen valores más cercanos a comparación de los de A (véase en la Figura 4.2) y de hecho la moda del histograma se encuentra entre el valor de ambos estados estacionarios.

5. EFECTO DEL RUIDO BIOQUÍMICO EN LAS DINÁMICAS DEL CIRCUITO

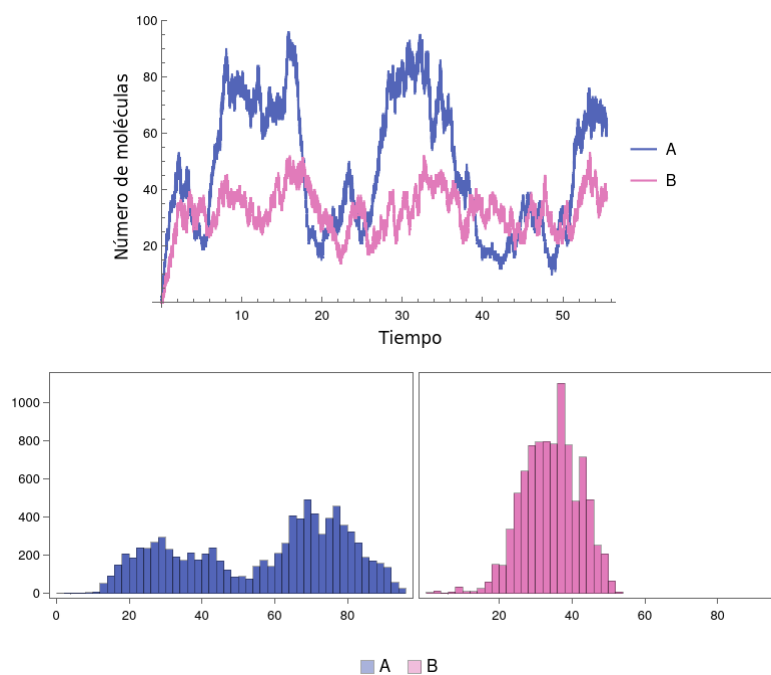


Figura 5.4: *Switch* biestable - Simulación estocástica del circuito con parámetros: $K_1 = 80$, $\alpha = 0.5$, $K_A = 40$, $K_B = 30$, $n_A = 8$, $n_B = 6$, $K_2 = 1$, $K_3 = 100$, $\beta = 0.25$, $K = 20$, $n = 3$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$ (figura superior). Histogramas del número de moléculas de ambas proteínas (figura inferior).

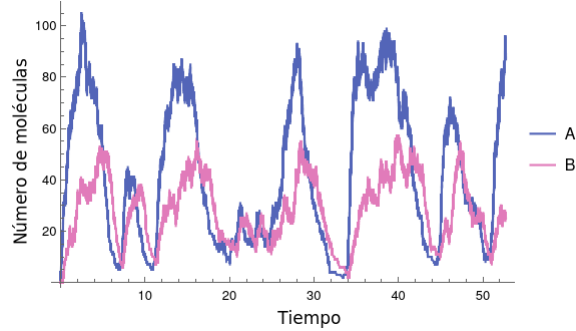


Figura 5.5: Simulación estocástica del Fenotipo (0,1,2) - Simulación estocástica del circuito con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 1$, $K_B = 0.7$, $n_A = 4$, $n_B = 4$, $K_2 = 1$, $K_3 = 40$, $\beta = 0.01$, $K = 20$, $n = 4$, $K_4 = 1$; con condiciones iniciales $A = 0$ y $B = 0$.

5.4. Sistema excitable inducido por ruido

Al introducir ruido al ejemplo de la región fenotípica (0,1,2) de la Figura 4.5 se obtiene una dinámica distinta a la de simplemente llegar a un estado estacionario. En la Figura 5.5, las proteínas parecen permanecer en valores $A \sim 80$ y $B \sim 40$ (correspondientes al punto de equilibrio estable) pero de vez en cuando hay fluctuaciones en su expresión que llegan a valores muy bajos, más de los que se esperarían de las fluctuaciones sobre un estado estacionario estable. Estas fluctuaciones coinciden con las trayectorias observadas en el mapa fase de la Figura 4.5; el sistema es atraído por el punto estable, pero es empujado fuera de él y da un recorrido por el mapa hasta regresar al nodo atractor.

5.5. Discusión

Las oscilaciones del modelo son robustas ante las fluctuaciones del ruido bioquímico, pues conservan amplitudes y periodos similares, aunque las amplitudes tienden a ser ampliadas gracias a las fluctuaciones y los periodos a ser más inexactos. El circuito además logra recapitular las oscilaciones inducidas por ruido que otros grupos han caracterizado ((18), (19), (38), (23)) al presentar dinámicas oscilatorias en las simulaciones estocásticas cuando el análisis de estabilidad y las simulaciones deterministas indican que el sistema llega a un estado estacionario. De acuerdo con lo que esos grupos han observado, las oscilaciones inducidas por ruido en este circuito también tienden a conservar periodos y amplitudes similares a las de oscilaciones deterministas que se encuentran cerca del punto de bifurcación.

Como se esperaba, los conjuntos de parámetros biestables pueden generar *state swit-*

5. EFECTO DEL RUIDO BIOQUÍMICO EN LAS DINÁMICAS DEL CIRCUITO

ching ante la presencia de ruido bioquímico. En el ejemplo de la Figura 5.4, solo una de las proteínas logra recapitular claramente las características del *state switching*. Aunque se ha observado que en otros ejemplos, es decir con otros parámetros, ambas proteínas presentan el *switching* y la bimodalidad, resulta interesante que esto no suceda en todos los casos.

El comportamiento de la Figura 5.5 se ha estudiado en otros circuitos sintéticos y biológicos (33) (5); en particular en la entrada al estado de competencia de subpoblaciones en poblaciones de *Bacillus subtilis*. Estas bacterias permanecen la mayor parte del tiempo fuera del estado competente, pero se ha propuesto que un sistema excitable es el responsable de llevarlos a este estado gracias a perturbaciones inducidas por ruido bioquímico. La topología base para este comportamiento es una combinación de retroalimentación positiva y negativa (7) como la de nuestro circuito.

Dinámicas evolutivas del circuito ante un ambiente fluctuante

6.1. Las poblaciones logran adaptarse ante ambientes fluctuantes

Como se explicó en el capítulo de Métodos, con el objetivo de caracterizar las dinámicas evolutivas de todos los elementos del circuito se realizaron dos tipos de experimentos: con la proteína A o la proteína B bajo selección. Simulaciones de cada uno de ellos se hicieron sobre diferentes genotipos iniciales, pasos mutacionales (M) y frecuencias de fluctuación ambiental (f). Las poblaciones se adaptaron mejor a algunas condiciones que a otras. En la Figura 6.1 se muestran *heatmaps* de los valores de adaptación máxima del promedio poblacional en los 10 últimos ciclos ambientales. En ambos tipos de experimentos y para todos los genotipos iniciales, los valores de adaptación más bajos se encuentran en la condición $f = 0.5$, es decir cuando la fluctuación ambiental se da cada 2 generaciones. Para las simulaciones con la presión selectiva sobre la proteína A también se observan valores de w más bajos para $M = 1.1$ en $f = 0.1$ y $f = 0.2$, excepto para el genotipo inicial 4. Un patrón similar se observa en los genotipos 1 y 3 de las simulaciones que seleccionan sobre la proteína B.

En las Figura 6.2 se muestran dos casos en los que las poblaciones lograron adaptarse a su ambiente fluctuante. En la izquierda de la figura se encuentran los valores de adaptación promedio (w) para el primer y último ciclo ambiental; en la derecha la distribución de los valores de la proteína que está bajo presión selectiva. En la parte superior de la Figura 6.2 están los resultados de la selección sobre A de la simulación con genotipo inicial 3, fluctuaciones ambientales $f = 0.02$ (cuando el ambiente cambia cada 50 generaciones) y paso mutacional $M = 3$. Al inicio del primer ciclo la población no se encuentra adaptada a su ambiente pero su valor de adaptación promedio comienza a

6. DINÁMICAS EVOLUTIVAS DEL CIRCUITO ANTE UN AMBIENTE FLUCTUANTE

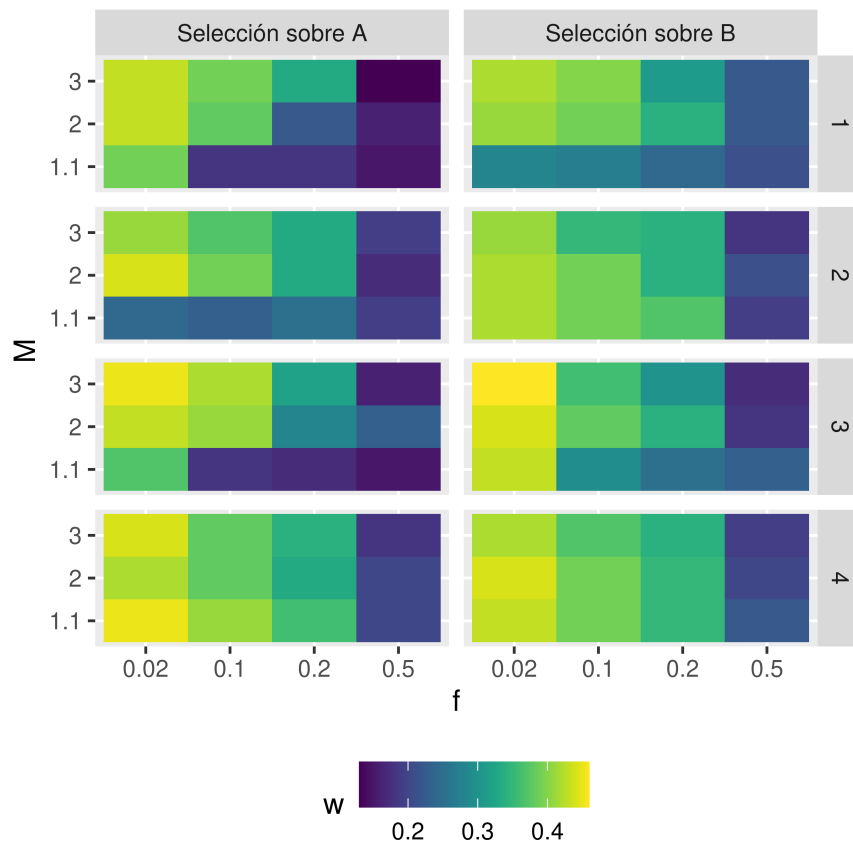


Figura 6.1: Adaptación promedio de las poblaciones - Valores de adaptación máxima del promedio de las poblaciones en los últimos 10 ciclos para cada experimento (grupos horizontales) y sus respectivos genotipos iniciales (grupos verticales) bajo las diferentes condiciones simuladas variando los parámetros M (filas) y f (columnas). Los genotipos iniciales son los siguientes: (1) estado estacionario alto en $P \sim 80$, (2) estado estacionario bajo en $P \sim 20$, (3) oscilaciones con picos y valles en $P \sim 80$ y $P \sim 20$, respectivamente y (4) biestabilidad con estados estacionarios en $P \sim 80$ y $P \sim 20$. Para las simulaciones que seleccionan sobre A, P corresponde a A, mientras que para las simulaciones que seleccionan sobre B, P corresponde a B. Los valores de adaptación se obtuvieron al promediar el valor máximo de adaptación promedio de la población en el ambiente bajo con el valor máximo de adaptación promedio de la población en el ambiente bajo, ambos dentro de los últimos 10 ciclos ambientales.

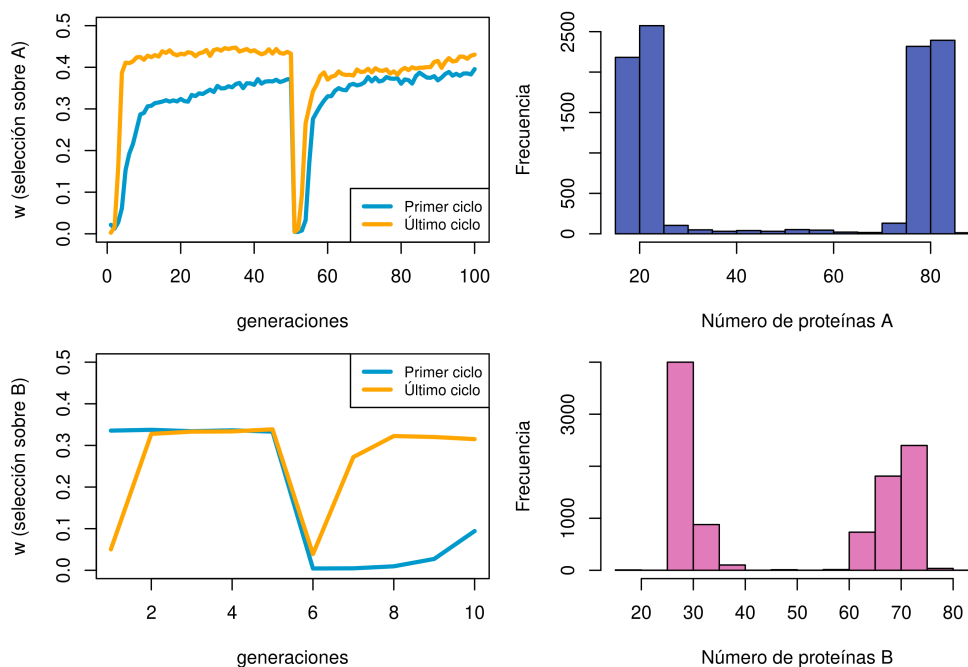


Figura 6.2: Casos en los que las poblaciones logran adaptarse - Promedio poblacional de adaptación (w) del primer y último ciclo ambiental; la población inicia en el ambiente bajo y después de $1/f$ generaciones cambia al ambiente alto. Histogramas de los promedios del número de proteínas en las 10,000 generaciones simuladas. Caso de selección sobre A con genotipo inicial 3, $f = 0.02$ y $M = 3$ (superior). Caso de selección sobre B con genotipo inicial 2, $f = 0.2$ y $M = 2$ (inferior).

subir en tan solo unas cuantas generaciones a valores considerablemente altos ($w \sim 0.3$); en la generación 50, cuando termina la primera época, el valor de adaptación decrece pero unas pocas generaciones después vuelve a subir. En el último ciclo la dinámica es similar, pero ahora la población alcanza valores promedio de adaptación más altos ($w \sim 0.4$) y el número de generaciones que pasan para que la población incremente w después del cambio ambiental es menor a comparación de la primer generación. La distribución de la proteína A es bimodal y las modas corresponden a los valores óptimos.

En la parte inferior de la Figura 6.2 están los resultados de la selección sobre B con genotipo inicial 2, fluctuaciones ambientales $f = 0.2$ (cuando el ambiente cambia cada 5 generaciones) y paso mutacional $M = 2$. Aquí, en la primera época la población ya se encuentra adaptada para el ambiente bajo pues el genotipo 2 tiene su estado estacionario alrededor de $B \sim 20$. En la generación 5 el valor de adaptación decrece y empieza a aumentar hasta las últimas generaciones del ciclo. Para el último ciclo ambiental, la población tarda menos generaciones después del cambio ambiental para

aumentar su adaptación promedio. La distribución de la proteína B también es bimodal pero sus modas no se encuentran exactamente en los valores óptimos aunque sí cerca. Los ejemplos escogidos tienen valores promedio de adaptación altos ($w > 0.3$) de acuerdo a la Figura 6.1, aunque el elegido de los experimentos de selección sobre A tienen un valor w más alto, en congruencia con las diferencias en las distribuciones y los valores de w en el último ciclo ambiental mostrado en la Figura 6.2; además de que las frecuencias ambientales son muy distintas y podrían influir en la diferencia de adaptación observada.

6.2. Estrategias para lidiar con ambientes fluctuantes

Se realizó el análisis de estabilidad para los genotipos de las últimas generaciones en el ambiente alto y en el ambiente bajo de cada linaje que sobrevivió un ciclo ambiental (según se explica en la sección de Métodos). Basándose en los resultados del análisis de estabilidad y en el número de mutaciones que ocurrieron a lo largo del linaje, se definieron las siguientes estrategias: (OSC) oscilaciones, el circuito oscila en todo el linaje y no ocurren mutaciones; (OSC GA) oscilaciones y adaptación genética, el circuito oscila al menos en los dos genotipos analizados y ocurren mutaciones; (BIS) biestabilidad, el circuito es biestable en todo el linaje y no ocurren mutaciones; (BIS GA) biestabilidad y adaptación genética, el circuito es biestable al menos en los dos genotipos analizados y ocurren mutaciones; (GA) adaptación genética, hay mutaciones pero estas causan que el circuito no permanezca en la misma dinámica (e.g. de oscilaciones a biestabilidad) o permanece en la misma dinámica pero esta no es biestable ni oscilatoria (e.g. monoestable). Finalmente, se clasificaron en (OTRO) a aquellos linajes en donde no ocurrieron mutaciones pero que la dinámica del genotipo no se identificó como oscilaciones ni biestabilidad (e.g. monoestabilidad).

En la Figura 6.3 se muestran las estrategias utilizadas en los linajes de las simulaciones de la sección anterior. Algunas tendencias generales fueron encontradas. La adaptación genética (GA) tiende a ser la estrategia dominante cuando el ambiente fluctúa a una frecuencia baja ($f = 0.02$) y cuando los pasos mutacionales son altos ($M = 2$ y $M = 3$), independientemente de la proteína que esté bajo selección y del genotipo inicial. También se observó que las estrategias (OSC GA) y (BIS GA) tienden a complementar a las oscilaciones (OSC) y a la biestabilidad (BIS) respectivamente, es decir que suelen dominar bajo las mismas condiciones.

También se encontraron otras tendencias pero estas no se mantienen en todos los casos. Por ejemplo, (OSC) es la estrategia más presente cuando el paso mutacional es bajo ($M = 1.1$) y cuando las fluctuaciones ambientales son muy frecuentes ($f = 0.5$); la excepción de este comportamiento son el caso en el que B está bajo presión selectiva y su genotipo inicial es el 2 y los casos con genotipo inicial 4 (i.e. biestabilidad), tanto para cuando A está bajo presión selectiva como para cuando B lo está. De manera

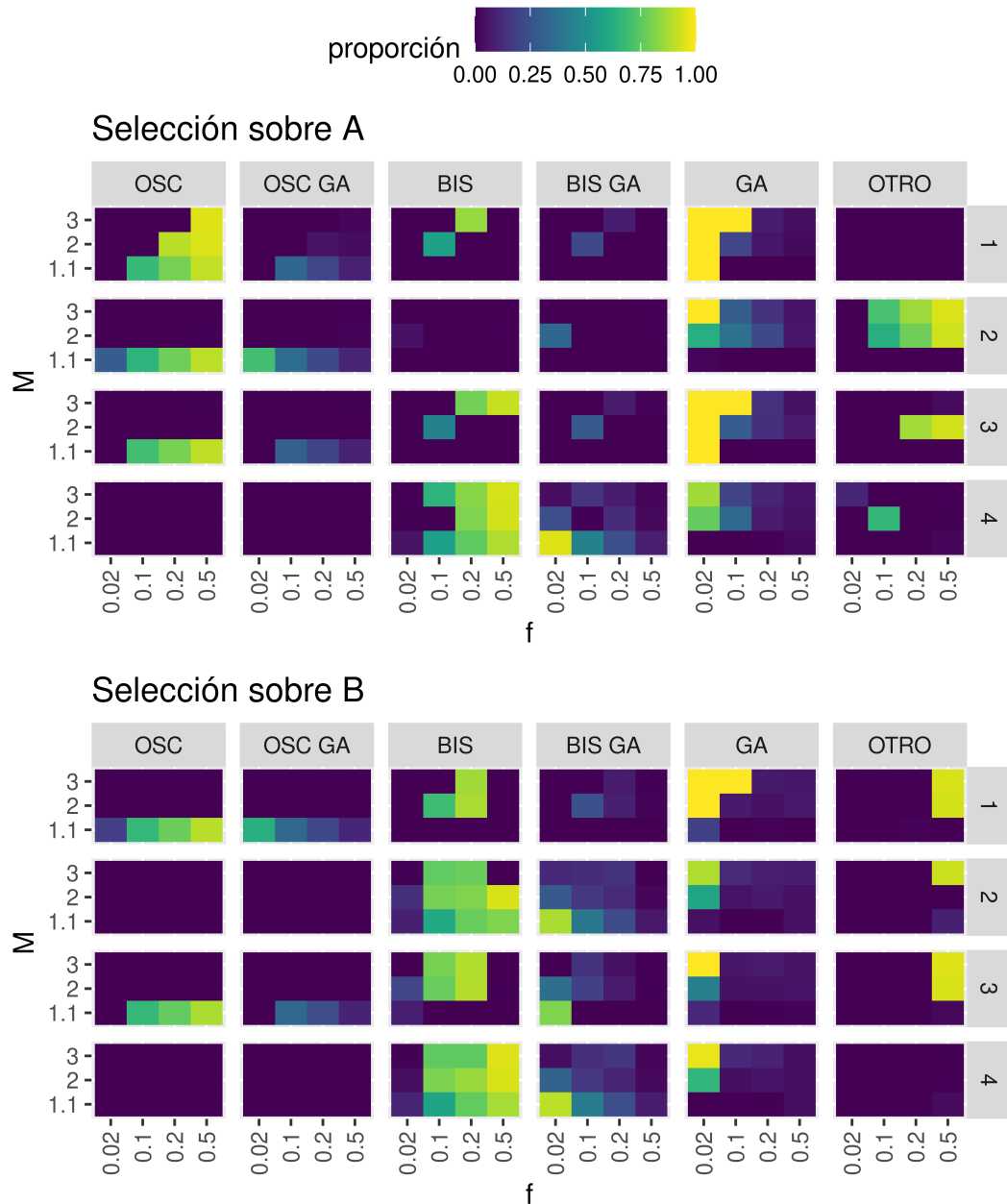


Figura 6.3: Estrategias de los linajes - Proporción de cada estrategia utilizada bajo las diferentes condiciones simuladas variando los parámetros M (filas) y f (columnas); agrupados según las estrategias descritas en el texto principal y los genotipos iniciales.

interesante, los linajes de los casos con genotipo inicial biestable no presentan las estrategias (OSC) ni (OSC GA).

Adicionalmente, en los experimentos que seleccionan sobre B y para los genotipos iniciales del 1 al 3, las estrategias que no fueron identificadas en ninguno de los casos anteriores (OTRO) dominan solo con la frecuencia $f = 0.5$ de cambio ambiental y en pasos mutacionales altos: $M = 2$ o $M = 3$; por otra parte, en los experimentos con A bajo presión selectiva su presencia es más variable entre los diferentes genotipos iniciales. Es importante remarcar que las estrategias de esta categoría requieren de más caracterización pues pudieran realmente pertenecer en otros grupos (véase más adelante). Finalmente, el patrón de (BIS) es menos claro pero parece ausente en los ambientes que fluctúan con menor frecuencia ($f = 0.02$) y junto a (BIS GA) es la estrategia dominante en casos en donde la población inicia con parámetros que generan biestabilidad (genotipo 4).

Aunque las oscilaciones dominan sobre otras estrategias en varias condiciones ambientales, como se puede observar en la Figura 6.1, estas condiciones suelen presentar un valor de adaptación menor en comparación con otros casos. Asimismo, las estrategias clasificadas en (OTRO) suelen también presentar baja adaptación en comparación. Por otra parte, en los genotipos iniciales del 1 al 3 los valores de adaptación más altos tienden a ocurrir con frecuencias de fluctuaciones ambientales bajas ($f = 0.02$) y pasos mutacionales altos ($M = 2$ y $M = 3$), en donde la adaptación genética es la estrategia preferida.

Con el objetivo de investigar más a fondo el efecto que utilizar diferentes estrategias tiene sobre la adaptación de la población, y como las distintas dinámicas emergentes (i.e. biestabilidad y oscilaciones) del circuito y la adaptación genética pueden ser implementadas como estrategias para lidiar con las fluctuaciones ambientales, examinamos el valor de adaptación promedio de los linajes (w_L) en tres casos en los que diferentes estrategias con dinámicas emergentes co-dominan y en un caso que utiliza solo adaptación genética; además, para cada uno de ellos se muestra una simulación estocástica de la dinámica de expresión que el o los genotipos de uno de los linajes presentan.

6.2.1. Dinámicas emergentes del circuito como estrategias adaptativas

6.2.1.1. Oscilaciones

En el caso en el que B se encuentra bajo presión selectiva, con genotipo inicial 1, $f = 0.02$ y $M = 1.1$ los linajes utilizan tres estrategias: (OSC), (OSC GA) y (GA). En la Figura 6.4 (izquierda) se muestran la distribución de las estrategias de acuerdo al número de mutaciones en el linaje y a w_L . Aquí destaca que los linajes con mayor valor de adaptación w_L son aquellos que utilizan las estrategias (OSC) y (OSC GA); aunque los linajes con estas estrategias llegan a tener valores similares a aquellos que tienen los linajes que utilizan (GA), es decir que no tienen exclusivamente valores altos.

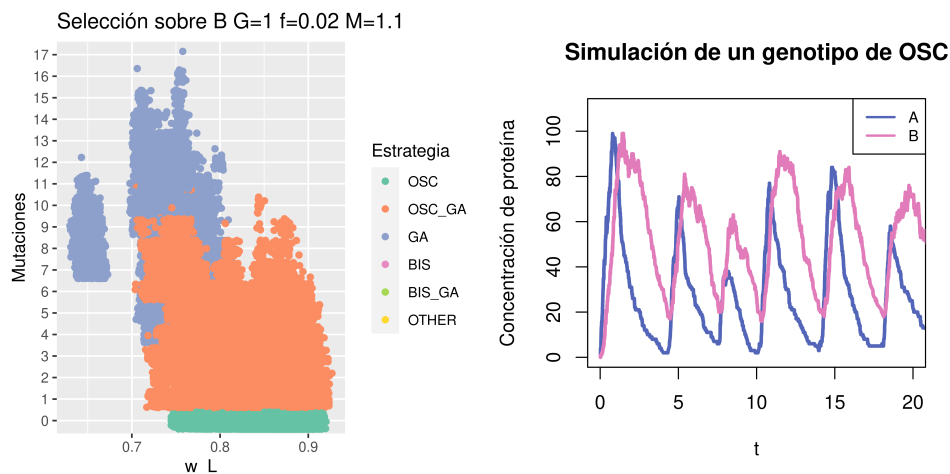


Figura 6.4: Oscilaciones como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo de uno de los linajes que utiliza OSC como estrategia, con parámetros: $K_1 = 160.08$, $\alpha = 0.99$, $K_A = 5.26$, $K_B = 16.61$, $n_A = 7.02$, $n_B = 15.38$, $K_2 = 1$, $K_3 = 160.24$, $\beta = 0.006$, $K = 37.43$, $n = 1.77$, $K_4 = 1$; se muestran dos decimales (derecha).

El valor de adaptación promedio de la población que se muestra en la Figura 6.1 es de $w = 0.28$; este valor no es particularmente alto, pero el valor de sus linajes sí llega a serlo.

La dinámica de uno de los genotipos de los linajes que utilizan (OSC) como estrategia se encuentra en la Figura 6.4 (derecha); el linaje de este genotipo tiene un valor de adaptación $w_L = 0.85$. Los valores de B alcanzar alturas y valles de 80 y 20 en sus oscilaciones, y estas tienden a permanecer más tiempo arriba y subir rápidamente al llegar a su valor mínimo. Notablemente, el periodo de las oscilaciones es similar a 4 unidades de tiempo, coincidiendo con el valor del ciclo de vida de las células de la población.

6.2.1.2. Biestabilidad - *State Switching*

En el caso en el que B se encuentra bajo presión selectiva, con genotipo inicial 2, $f = 0.02$ y $M = 1.1$ (las mismas condiciones que en el caso anterior pero con un genotipo inicial distinto), los linajes utilizan tres estrategias: (BIS), (BIS GA) y (GA) (Figura 6.5, izquierda). A diferencia del caso de la sección anterior, los valores de w_L de algunas estrategias se encuentran más restringidos y el valor promedio de adaptación de la población es mayor $w = 0.42$ (véase Figura 6.1), la posible relación de estos dos

6. DINÁMICAS EVOLUTIVAS DEL CIRCUITO ANTE UN AMBIENTE FLUCTUANTE

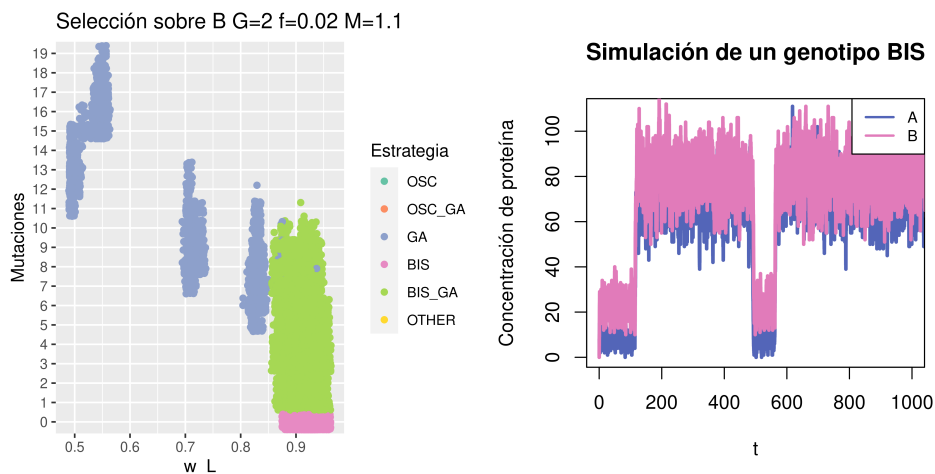


Figura 6.5: Biestabilidad-*State Switching* como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo de uno de los linajes que utiliza BIS como estrategia, con parámetros: $K_1 = 70.8$, $\alpha = 0.36$, $K_A = 18.14$, $K_B = 17.58$, $n_A = 11.91$, $n_B = 6.04$, $K_2 = 1$, $K_3 = 82.46$, $\beta = 0.14$, $K = 17.4$, $n = 2.23$, $K_4 = 1$, se muestran dos decimales (derecha).

factores se considera en la sección de Discusión. Aquí los linajes que utilizan (BIS) y (BIS GA) como estrategia parecen presentar menor variación en sus valores de w_L , siendo que estos están concentrados en valores de $w_L > 0.85$, mientras que los linajes que utilizan (GA) presentan valores de w concentrados en $w_L < 0.85$.

Uno de los genotipos de los linajes que utilizan (BIS) se encuentra en la derecha de la Figura 6.5; su linaje tiene un valor de adaptación $w_L = 0.92$. En la simulación se puede observar claramente que el efecto del ruido bioquímico permite que el genotipo biestable presente *switching* epigenético. Los linajes con este genotipo presentan *epimutaciones* que les permiten cambiar de un estado estacionario a otro. En la simulación se ve que ambas proteínas (de particular importancia la proteína B, ya que es la que está bajo selección), permanecen dando pequeñas fluctuaciones cerca de uno de los estados, para después de una mayor fluctuación, cambiar al otro.

6.2.1.3. Oscilaciones inducidas por ruido

El último caso sin mutaciones analizado es aquel en el que la presión selectiva es sobre A, con genotipo inicial 2, $f = 0.5$ y $M = 3$. Las estrategias que utilizan estos linajes son principalmente (GA) y las clasificadas como (OTRO), aunque también se encuentran representadas (OSC GA) y (BIS GA); a diferencia de los dos casos anterior-

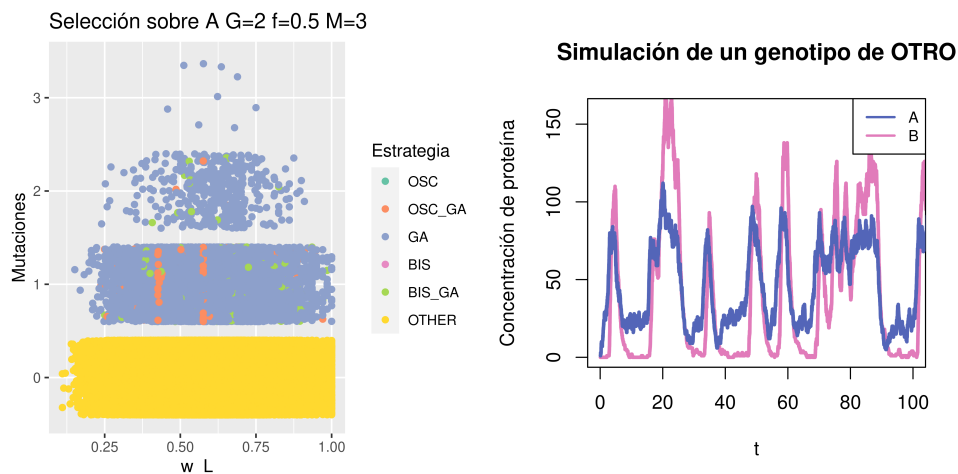


Figura 6.6: Oscilaciones inducidas por ruido como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo con parámetros: $K_1 = 100$, $\alpha = 0.25$, $K_A = 38.58$, $K_B = 8.69$, $n_A = 5.45$, $n_B = 1.07$, $K_2 = 1$, $K_3 = 150$, $\beta = 0.007$, $K = 69.45$, $n = 16$, $K_4 = 1$, se muestran dos decimales (derecha).

res, las estrategias principales presentan una variación considerable en sus valores de adaptación, tomando valores menores a 0.25 y hasta 1 (Figura 6.6, izquierda). Este caso es de los que presentan menor valor promedio de adaptación en la población ($w = 0.19$) (véase Figura 6.1).

Uno de los genotipos que utiliza la estrategia clasificada como (OTRO) se encuentra en la Figura 6.6 (derecha) y el linaje al que pertenece tiene una adaptación de $w_L = 0.70$. El análisis de estabilidad de este genotipo indica que el sistema tiene un único estado estacionario estable, por lo que de manera determinista no presenta oscilaciones; por lo tanto, estas son oscilaciones inducidas por ruido. De acuerdo a la simulación, la proteína bajo selección (A), alcanza picos y valles cerca de los valores óptimos 80 y 20. A diferencia de las oscilaciones observadas en la derecha de la Figura 6.4, estas presentan un periodo inconsistente, permaneciendo tiempos irregulares en valores altos y bajos de A.

Particularmente para este caso más del 99% de los genotipos de los linajes que utilizan estrategias clasificadas como (OTRO) son monoestables. Es importante resaltar que esto no quiere decir que todas las estrategias clasificadas como (OTRO) correspondan a esta dinámica, pues análisis de otros casos resultaron en que algunos linajes realmente pertenecían a la clasificación de (BIS) pero nuestra manera de analizarlos no logró

6. DINÁMICAS EVOLUTIVAS DEL CIRCUITO ANTE UN AMBIENTE FLUCTUANTE

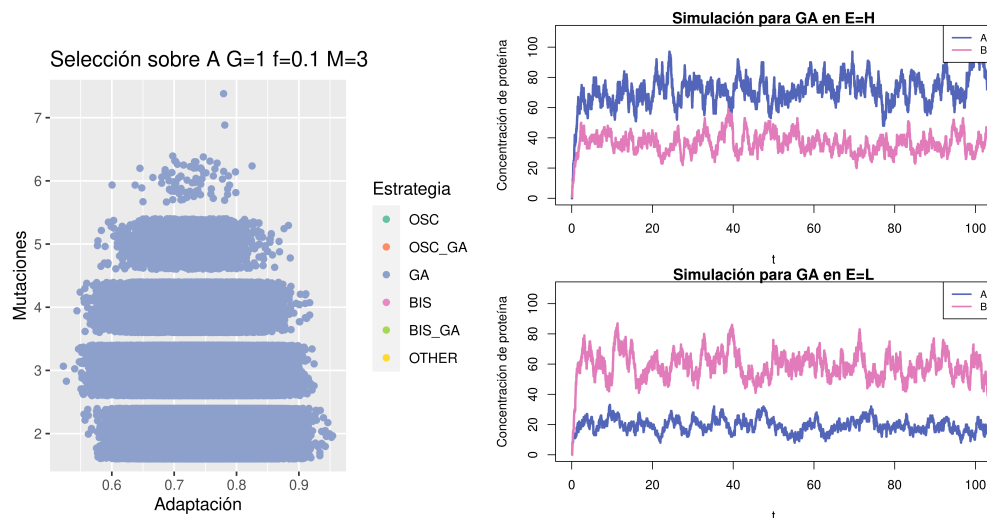


Figura 6.7: Adaptación genética como estrategia adaptativa - Distribución de las estrategias de acuerdo al número de mutaciones en el linaje y al valor promedio de adaptación en el linaje (w_L) (izquierda). Simulación estocástica del genotipo del ambiente E=H con parámetros: $K_1 = 318.35$, $\alpha = 0.06$, $K_A = 0.01$, $K_B = 0.03$, $n_A = 0.03$, $n_B = 0.2$, $K_2 = 1$, $K_3 = 113.41$, $\beta = 0.035$, $K = 17.93$, $n = 0.12$, $K_4 = 1$. Simulación estocástica del genotipo del ambiente E=L con parámetros: $K_1 = 439.61$, $\alpha = 0.03$, $K_A = 0.01$, $K_B = 0.09$, $n_A = 0.02$, $n_B = 0.5$, $K_2 = 1$, $K_3 = 113.41$, $\beta = 0.035$, $K = 17.93$, $n = 0.12$, $K_4 = 1$, se muestran dos decimales (derecha).

clasificarlos como tal.

6.2.2. Adaptación genética como estrategia adaptativa

La estrategia dominante en el caso en donde se selecciona sobre A, con genotipo inicial 1, $M = 3$ y $f = 0.1$ la estrategia dominante es (GA) (Figure 6.7, izquierda). Al igual que se ha visto en casos anteriores, tiende a tener una amplia variación de w_L . Para estos linajes, el valor promedio de su población es $w = 0.38$ (véase Figura 6.1). En la derecha de la Figure 6.7 se muestran los dos genotipos de la última generación del linaje al final del la época ambiental en el ciclo. El valor promedio de adaptación del linaje es $w_L = 0.82$. El genotipo del ambiente alto tiene valores de $A \sim 80$, mientras que el genotipo de ambiente bajo valores de $A \sim 20$. El análisis de estabilidad no logró encontrar el estado estacionario para estos genotipos, pero simulaciones deterministas con condiciones iniciales $A = 0$ y $B = 0$ muestran que el sistema se va a un estado estacionario, por lo que podría ser monoestable. Los parámetros que cambian más significativamente entre los dos genotipos son los de la ecuación diferencial de A (ecuación

(3.24)).

6.3. Discusión

Las poblaciones utilizaron diversas estrategias para lidiar con las fluctuaciones ambientales a las que fueron sometidas. Bajo ciertas condiciones se observó que la población alcanzaba mayores valores de adaptación que en otras, además de que ciertas estrategias estaban más representadas en ciertos valores de M y de f . En general, las poblaciones se adaptan pobremente ante fluctuaciones ambientales con frecuencia $f = 0.5$. Esto podría significar que ninguna de las estrategias exploradas aquí resulta adecuada para ambientes que fluctúan con frecuencias tan altas. Además, la tendencia en todas las simulaciones parece indicar que cuando el ambiente fluctúa con frecuencias bajas (e.g. $f = 0.02$) la adaptación promedio de la población es mayor, independientemente de la proteína que esté bajo selección. Para algunos genotipos iniciales se observó que la adaptación era baja cuando el paso mutacional era $M = 1.1$; esto podría deberse a que dichos genotipos se encontraban muy lejos en el espacio de parámetros de una solución que ofreciera valores de adaptación mayores, por lo que la población se quedaban atrapada en soluciones que podrían pensarse como mínimos locales.

Al analizar las estrategias utilizadas por los linajes en diferentes condiciones ambientales logramos recapitular resultados de otros trabajos en donde la adaptación genética se beneficia en pasos mutacionales altos y cambios ambientales de baja frecuencia (16); En Gómez-Schiavon et al. (16) también observaron que el *switching* epigenético se veía beneficiado cuando las frecuencias de las fluctuaciones ambientales son altas (aquí asumimos que (BIS) utiliza *switching* epigenético, véase más adelante). Este comportamiento claramente se recapitula con el genotipo inicial 4 de cada una de las dos proteínas bajo selección, y aunque el patrón no es obvio en otros genotipos iniciales, (BIS) y (BIS GA) se presentan en condiciones con $f < 0.02$. Es importante mencionar que en dicho trabajo las frecuencias más alta, es decir el valor de f más grande, corresponde a $f = 0.1$, mientras que aquí los valores de las fluctuaciones ambientales correspondieron a $f = \{0.02, 0.1, 0.2, 0.5\}$; para investigar si este patrón también se recapitula para otros genotipos iniciales, más simulaciones tendrían que hacerse en el rango $f \in [0.02, 0.1]$.

Los dos tipos de experimentos (selección sobre A y sobre B) permitieron explorar de una manera más completa el circuito; esta exploración no logró identificar una diferencia clara entre seleccionar una proteína sobre otra. Por otra parte, explorar las mismas condiciones desde diferentes genotipos iniciales permitió observar que las poblaciones no necesitan iniciar en la región de parámetros de biestabilidad u oscilaciones para llegar a utilizar estas estrategias, pues incluso con pasos mutacionales bajos ($M = 1.1$), las poblaciones lograron llegar a dichas regiones del espacio de parámetros. Esto también permitió observar que en ciertas condiciones algunas estrategias logran dominar en

6. DINÁMICAS EVOLUTIVAS DEL CIRCUITO ANTE UN AMBIENTE FLUCTUANTE

los linajes independientemente del genotipo inicial (e.g. la adaptación genética tiende a prevalecer en ambientes que fluctúan con baja frecuencia y pasos mutacionales altos). Por otra parte, algunas estrategias solo se presentaron con ciertos genotipos iniciales; por ejemplo, las oscilaciones tienden a seleccionarse en pasos mutacionales bajos ($M = 1.1$) excepto con el genotipo inicial 2 del experimento con B bajo presión selectiva y en los genotipos iniciales 4 para ambos tipos de experimentos (A y B bajo presión selectiva). De estos dos últimos casos, resulta interesante que cuando el genotipo inicial era biestable, no se presentaba la estrategia (OSC) en los linajes. Una posible explicación es que (OSC) ofreciera menor valor de adaptación que (BIS) y (BIS GA), y al toda la población iniciar con estos genotipos, si un linaje celular llegaban a presentar un fenotipo oscilatorio, este se extinguía al competir con las otras estrategias. Esto se sospecha principalmente porque en las simulaciones en las que (OSC) tendía a ser la estrategia dominante, el w poblacional era bajo. En general, explorar diferentes genotipos iniciales resultó ser informativo, aunque ya que estas son simulaciones estocásticas más replicas son necesarias pues desde un mismo genotipo diferentes patrones de estrategias podrían surgir.

Para investigar cómo la elección de ciertas estrategias influye en la adaptación de la población se analizaron las estrategias de los linajes en relación con el promedio de su valor de adaptación en todo el linaje durante el ciclo ambiental (w_L). Esto permitió observar que aunque la población tenga valores de adaptación promedio no tan altos ($w < 3$), sus linajes pueden alcanzar valores altos de w_L . Además, en los casos considerados (secciones 6.2.1-6.2.2) se observaron distintas distribuciones en los valores de w_L , donde las estrategias (BIS) y (BIS GA) del caso de biestabilidad (Figura 6.5) tienden a tener una distribución restringida a valores altos de w_L . Esto puede ser relevante porque de los casos analizados, este presenta el valor de adaptación poblacional mayor. En cambio, el valor de adaptación poblacional menor sucede en el caso de la sección 6.2.1.3, en donde los linajes presentan la mayor variación de w_L .

Finalmente, cuatro estrategias utilizadas por los linajes fueron analizadas: oscilaciones, biestabilidad (*state switching*), oscilaciones inducidas por ruido y adaptación genética. Las oscilaciones llegaban a presentar valores altos de w_L . El caso analizado parecía tener el mismo periodo que la duración de un ciclo de vida de los miembros de la población, mientras que la fluctuación ambiental sucedía cada 50 generaciones. Lo que podría estar sucediendo es que cuando la población se encuentra en el ambiente bajo, el final de una generación coincide con el valle de la oscilación, obteniendo así valores de adaptación altos. A lo largo de las generaciones, gracias a que el periodo no es preciso por efectos del ruido bioquímico, los picos y valles de las oscilaciones se desfasan, causando así que cuando termina la época y sucede el cambio de ambiente, ahora los picos coinciden con el cambio de generación. Algo a considerar es que en los ejemplos de oscilaciones presentados este trabajo, se ha observado que estas no son simétricas. Es decir que permanecen un tiempo distinto arriba y abajo. En un trabajo futuro sería interesante ver como se comportan las poblaciones con una presión selectiva asimétrica

en los ambientes.

El *state switching* y la adaptación genética lograron valores altos en el promedio de adaptación de las poblaciones. Para el primero, el rol del ruido bioquímico parece crucial, pues gracias a las fluctuaciones los linajes pudieron presentar valores w_L altos al adaptarse a ambos ambientes; es probable que el cambio de estado coincidiera con los cambios ambientales. Hasta ahora hemos asumido que los genotipos biestables están siendo seleccionados por presentar *state switching*, pero en realidad solo hemos estudiado un ejemplo de esta estrategia y una mayor caracterización es necesaria; simulaciones deterministas del modelo evolutivo podrían resultar útiles, siendo que en ellas no habría cambio de estado por falta de fluctuaciones y se mantendrían siempre sobre uno de los estados, la estrategia (BIS) pudiera no conferir ventaja en estas condiciones si es verdad que se selecciona por su capacidad de generar *switching* epigenético. En cuanto a la adaptación genética, los pasos mutacionales grandes pueden permitirle pasar del genotipo óptimo de un ambiente al genotipo óptimo del otro en pocas mutaciones. Por la manera de mutar de nuestro circuito, cuando una mutación ocurre solo algunos parámetros cambian; en un trabajo futuro, estudiar los parámetros que tienden a cambiar para lograr la adaptación resultaría interesante.

El rol del ruido bioquímico también permitió que las oscilaciones inducidas por ruido llegaran a presentar valores w_L altos. Aquí las oscilaciones son muy irregulares en su periodo y podrían resultar en una estrategia capaz de competir con el *state switching* cuando el cambio ambiental ocurre a frecuencias tan altas como $f = 0.5$, pues los picos de las oscilaciones pudieran ocurrir con mayor frecuencia que el cambio causado por la *epimutación*. En el futuro, caracterizar las frecuencias posibles de las *epimutaciones* en este circuito también pudiera resultar informativo sobre por qué las poblaciones con $f = 0.5$ y en donde (BIS) es la estrategia dominante presentan valores de adaptación poblacional promedio bajos.

Conclusiones y Perspectivas

En este trabajo se estudiaron las dinámicas de expresión de un circuito de regulación génica con retroalimentación positiva y negativa y las dinámicas evolutivas de poblaciones con dicho circuito sometidas a ambientes fluctuantes. Observamos que el sistema es capaz de generar oscilaciones, biestabilidad o alcanzar un único estado estacionario estable (i.e. monoestabilidad). En presencia del ruido bioquímico, además puede presentar *switching* epigenético, generar oscilaciones inducidas por ruido o comportarse como un sistema excitable inducido por ruido. Nuestro modelo de dos dimensiones puede presentar un repertorio amplio de dinámicas que resultan como potenciales fenotipos del sistema, siendo el papel del ruido bioquímico esencial para algunas de ellas. Si bien todas estas dinámicas ya habían sido identificadas en distintos trabajos (e.g. (16), (23), (33)), bajo nuestro conocimiento no se habían reportado en un mismo sistema.

A partir de nuestros resultados se puede concluir que un análisis determinista de las dinámicas de expresión resulta insuficiente para caracterizar el verdadero potencial fenotípico de nuestro circuito de regulación, sugiriendo que pudiera no ser la mejor práctica ignorar el papel del ruido bioquímico en otros circuitos de regulación. Sería interesante ver si las dinámicas que encontramos en este modelo se mantienen al añadirle más proteínas al circuito (i.e. más dimensiones), mientras se mantiene la estructura del circuito (i.e. retroalimentación positiva y negativa), o si surgen nuevas dinámicas inducidas por ruido. Al aumentar las dimensiones del sistema además pueden surgir otras dinámicas en las simulaciones deterministas, como caos, que también pudieran ser de interés como estrategias evolutivas y revelar condiciones en las que ese tipo de dinámicas son funcionales para las células.

Los resultados de las simulaciones evolutivas indicaron que las dinámicas que surgen como propiedades emergentes del circuito (las oscilaciones) y las emergentes del circuito a consecuencia del efecto del ruido bioquímico (*switching* epigenético y oscilaciones inducidas por ruido) son capaces de surgir y permanecer en las poblaciones como estrategias para lidiar con ambientes fluctuantes sin necesidad de mutaciones, siendo capaces de (en ciertos ambientes) competir exitosamente con la adaptación genética.

7. CONCLUSIONES Y PERSPECTIVAS

Bajo nuestro conocimiento, no se había explorado el potencial de los osciladores bioquímicos ni de las oscilaciones inducidas por ruido como estrategias adaptativas para lidiar con ambientes fluctuantes. Adicionalmente, si bien ya se había especulado que las oscilaciones inducidas por ruido podrían tener una función en relojes circadianos al proporcionarle robustez al sistema (23), no se había explorado una funcionalidad como la que presentamos en este trabajo. Los resultados de este trabajo apoyan la idea de que las dinámicas inducidas por ruido pueden tener funciones en los sistemas biológicos. Si bien en esta primera exploración no se identificaron a los sistemas excitables inducidos por ruido bioquímico como estrategias adaptativas, no descartamos que éstos pudieran tener una función adaptativa en otras condiciones evolutivas a ser detectada en una exploración más exhaustiva, como evaluando fluctuaciones aleatorias o ambientes con presión selectiva asimétrica. Estas últimas en particular podrían favorecerla, pues el sistema excitable inducido por ruido tiende a pasar mayor tiempo cerca de su estado estacionario y los recorridos por el plano fase son espontáneos. Esta clase de ambiente también pudiera favorecer al tipo de oscilaciones que puede generar nuestro circuito, pues son asimétricas.

Aquí nos enfocamos en caracterizar las dinámicas evolutivas para 4 valores de f y 3 de M , desde cuatro genotipos iniciales y con presión selectiva sobre las dos proteínas que conforman el circuito. Más réplicas son necesarias para ganar certeza sobre nuestras observaciones, particularmente considerando que las simulaciones realizadas son estocásticas tanto en el modelo de expresión génica como en el modelo evolutivo. Además, probar más valores de M y f puede ayudar a generar una perspectiva más clara de las condiciones que favorecen cada una de las estrategias estudiadas. Por otra parte, realizar simulaciones con variaciones de otros parámetros del modelo evolutivo, como la tasa de mutación (u), el tamaño de la población (N) y la presión selectiva (s_T) también resultarían útiles para probar el alcance de nuestras conclusiones.

A.1. Análisis de estabilidad de los estados estacionarios

Lenguaje de programación: MATLAB

```
%Evaluate SS behavior of a 2 species model with positive
%and negative feedback loop.

%Input: Parameters
%Output: Returns calculated behavior.

%Results: behaviors= x.yzuv
%  x = number of stable points;
%  y = STABLE SPIRAL; number of stable points
%with negative eigenvalues and complex component
%  z = UNSTABLE SPIRAL; number of stable points
%with positive eigenvalues and complex component
%  u = number of detected intersections
%  v = indicates if there was a break to imaginary numbers
%(1); this is relevant because some intersections may not be
%detected for this reason.
%      Behaviors of the form x.yzuv1 should be interpreted
%with caution
%
% Possible Behavior List
% 0.011 Oscillations
% 1.001 Monostable with 1 intersection
% 1.003 Monostable with 3 intersection
% 1.101 Monostable spiral with 1 intersection
% 1.103 Monostable spiral with 3 intersection
% 2.003 Bistable with 3 intersection
```

A. CÓDIGOS

```
% 2.103 Bistable spiral with 3 intersection

% 1.013 Monostability with possible noise induced
%excitability
% 1.0131
% 0.0111

tic

%ODE system
syms alpha k1 ka kb na nb a b;
dA = symfun(k1*(alpha*(kb^nb/(kb^nb+b^nb))+
(1-alpha*(kb^nb/(kb^nb+b^nb)))*(a^na/(a^na+(ka^na/kb^nb)
*(b^nb+kb^nb))))-a,[alpha, k1, ka, kb, na, nb, a, b]);
syms beta k3 k n a b;
dB = symfun(k3*(beta+(1-beta)*(a^n/(a^n+k^n)))-b,
[beta, k3, k, n, a, b]);

%Jacobian Matrix
syms a_ss b_ss
func = [dA(alpha, k1, ka, kb, na, nb, a_ss, b_ss),
dB(beta, k3, k, n, a_ss, b_ss)];
var = [a_ss, b_ss];
jacob= jacobian(func,var);
eigenval_symb=eig(jacob);

eqs=[dA(alpha, k1, ka, kb, na, nb, a, b)==0,b>0];
nullclineA = solve(eqs,b,'Real',true,'ReturnConditions',true);
nullclineA = nullclineA.b;
nullclineB = solve(dB(beta, k3, k, n, a, b)==0, b);

interval=100;%Interval number for search

%Declare vectors for intersection search
dif=ones(2,interval+1);
difsmall=ones(2,interval+1);
dif_nonzero=ones(2,interval);
mult_vectors=ones(1,interval);

%Obtain parameter values (these values are examples)

k1 = 90;
alpha = 0.5;
```

```

ka = 30;
kb = 40;
na = 8.0;
nb = 5.0;
k3 = 150.0;
beta = 0.01;
k = 28;
n = 15;
k2=1;
k4=1;

nullclineA_s=subs(nullclineA);
nullclineB_s=subs(nullclineB);

%Find Fixed Points: Find change in sign (which occurs after
%intersection)

%Find Fixed Points: Find change in sign
dif(2,2:length(dif))=0.0001:((k1-0.0001)/(interval-1)):k1;

breaks=-1; %breaks starts in -1 because the final number
%evaluated (k1) is imaginary in
%c=nullclineA(k1)-nullclineB(k1)

for j=2:1:length(dif)
    c=subs(nullclineA_s, a, dif(2,j))-subs(nullclineB_s, a,
    dif(2,j)); %Evaluate nullclines difference in
    %a=dif(2,j)
        if (isreal(c)) %If difference is in reals,
            %asing difference
                dif(1,j)=c;
            else %If the difference does not exist in reals,
                %assign 0
                    breaks=breaks+1;
                    dif(1,j)=0;
            end
        end
end

%If the nullcline does not break to imaginary values,
%the value after the intersection is then further
%searched
if(~breaks) %Multiplication of vector considers
%firstvalue*1, since nullclineA<nullclineB

```

```
%gives negative result in dif(1,1) if intersection
mult_vectors=dif(1,1:(length(dif)-1)).
*dif(1,2:length(dif));
%Multiply vectors: negative result indicates
%change in sign (intersection)
index_a= find(mult_vectors < 0);
%Returns index where the value of the
%multiplication is negative
index_a=index_a+1; %add one to index because dif(1,1)
%is always 1 and not the difference
value_a=dif(2,index_a);
if(value_a)
    value_a_small=value_a;
    i=1;
    for u=1:1:length(index_a) %Evaluate all detected
%intersections, extensive search
        if (value_a(u)==dif(2,2)) %If the values is the
%first one we cannot compare and value_a
%stays the same
            i=i+1;
            continue
        end
        value_a2=value_a(u); %The comparison is between
%value_a and the previous value
        value_a1=dif(2,index_a(u)-1);
        difsmall(2,2:length(difsmall))=value_a1 :
((value_a2-value_a1)/(interval-1)):value_a2;
%Evaluate 100 points from value_a1 to value_a2
        for j=1:1:length(difsmall)
            c=subs(nullclineA_s, a, difsmall(2,j))-
            subs(nullclineB_s, a, difsmall(2,j));
            %Evaluate nullclines difference in
            %a=difsmall(2,j)
            difsmall(1,j)=c;
        end
        mult_vectors=difsmall(1,1:(length(difsmall)-1))
        .* difsmall(1,2:length(difsmall));
        %Multiply vectors: negative result indicates
        %change in sign (intersection)
        mult_vectors(1)=0;
        %Ignore first value (value_a1*1)
        index_a_final= find(mult_vectors < 0);
        %Returns index where the value of the
```

```

        %multiplication is negative
        index_a_final=index_a_final+1;
        if (length(index_a_final)>1)
            for w=1:length(index_a_final)
                value_a_small(i)=
                    difsmall(2,index_a_final(w));
                %Value of A is updated
                i=i+1;
            end
        else
            value_a_small(i)=difsmall(2,index_a_final);
            i=i+1;
        end
    end
    value_a=value_a_small;
end
else
breaks=1;
count_nonzero=1;
beforebreak=0;
for j=1:length(dif)
    if(dif(1,j))
        if(count_nonzero>1)
            %if this is not the first non-zero value
            if((dif(1,j)<0 && ~dif(1,j-1)) &&
                dif_nonzero(1,count_nonzero)>0)
                beforebreak=1;
            end
        end
        count_nonzero=count_nonzero+1;
        dif_nonzero(1,count_nonzero)=dif(1,j);
        dif_nonzero(2,count_nonzero)=dif(2,j);
    end
end
mult_vectors=dif_nonzero(1,1:(count_nonzero-1))
.*dif_nonzero(1,2:(count_nonzero));
index_a= find(mult_vectors<0);
if(~beforebreak)
    index_a=index_a+1;
end
%Returns index where the value of the multiplication
%is negative
value_a=dif_nonzero(2,index_a);

```

A. CÓDIGOS

```
end

if(dif(1,length(dif)-1)>0) %if the last value of
%c=nullclineA-nullclineB is positive , then the nullclines
%intersect but it is not detected
    value_a(length(value_a)+1)=dif(2,length(dif)-1);
    %add last positive value to intersections
end

%Evaluate Jacobian Matrix in value_a and
%count stable steady states
stable_points=0;
complex_neg=0;
complex_pos=0;
value_b=value_a;

if (value_a)
    for u=1:length(value_a)
        value_b(u)=subs(nullclineA_s , a , value_a(u));
        a_ss=value_a(u);
        b_ss=value_b(u);
        %jacobfp=subs(jacob);
        eigenval=subs(eigenval_symb);
        if(double(eigenval(1))<0 && double(eigenval(2))<0)
            %Determine stability
            %only considers stable points when both
            %eigenvalues are negativa (in this way saddle
            %points are correctly evaluated as unstable
            stable_points=stable_points+1;
            if(~isreal(eigenval)) %Determine if there is
            %complex component (negative eigen)
                complex_neg=complex_neg+1;
            end
        elseif(~isreal(eigenval))
            complex_pos=complex_pos+1;%Determine if
            %there is complex component (positive eign)
        end
    end
end

%Assign behavior
behavior=stable_points;
```

```

    if (complex_neg)
        behavior=behavior+complex_neg/10;
    end
    if (complex_pos)
        behavior=behavior+complex_pos/100;
    end
    if (breaks)
        behavior=behavior+breaks/10000;
    end

    behavior=behavior+length(value_a)/1000
toc

```

A.2. EvoDynamics_OscillatoryModel.cpp (Selección sobre B)

Lenguaje de programación: C++

```

// Evolutionary dynamics of oscillatory model in fluctuating
// environments
// SIMULATIONS
//
// Created by Mariana Gomez-Schiavon
// October 2015
/// Adapted for oscillatory model by Sophia Orozco
// October 2021

// Libraries:
#include <iostream>
#include <stdlib.h>
#include <fstream>
#include <stdio.h>
#include <algorithm>
#include <time.h>
#include <cmath>
#include <boost/array.hpp>
#include <boost/numeric/odeint.hpp>
#include "GeneNetwork_OscillatoryModel_B_unfpar.h"

using namespace std;
//2 Variable System

```


A. CÓDIGOS

```
typedef boost::array< double , 2 > state_type;
typedef boost::numeric::odeint::runge_kutta_cash_karp54
< state_type > rkck54; //Controlled Stepper
typedef boost::numeric::odeint::controlled_runge_kutta
< rkck54 > ctrl_rkck54;

int main(int argn, char *args[])
{
    //////////////////////////////////////
    // Simulation constants:
    #define TMAX 4 //Maximum time to simulate each cell during
    //one generation.
    const double tmax_deterministic = 4.0;
    int GMAX = 10000; //No. of generations to simulate.
    int Cp = 100; //No. of cycles to print lineages' info
    int Ep = 10; //No. of epoch to print individuals' info
    //only first 9 & last generation per epoch

    // Initial protein concentrations:
    int A0 = 20;
    int B0 = 10;
    // Initial environmental state:
    int E0 = 1; //0 - High; 1 - Low
    // Fixed parameters:
    double k2 = 1; // Protein degradation constant
    double k4 = 1; // Protein degradation constant
    // Evolving parameters (1 - Evolve; 0 - Fixed):
    parE PARE;
    PARE.k1 = 1; // Maximum synthesis rate
    PARE.alpha =1; // Basal synthesis
    PARE.nA = 1;
    PARE.nB = 1; // Hill coefficient
    PARE.kA = 1; // Affinity constant
    PARE.kB = 1; // Affinity constant
    PARE.beta =1;
    PARE.k3 = 1;
    PARE.n = 1;
    PARE.k = 1;
    // Optimal protein number:
    double BOPT[2] = {80,20}; // {E.High,E.Low}; Selection on B
    double V = 0.2; // Width measure for Lorentzian fitness
    // Replicas:
    int numRep = 1; // Total number of replicas to run.
```

```

// Seeds for the random number generator.
long seeds[10] = {-17,-23,-7,-3,-5,-9,-11,-13,-15,-19};
////////////////////////////////////
if (argn < 17)
    {
        cerr << "Error! Input 16 arguments: N; nu; sT;
u; M; Algorithm: (0) Deterministic ,
(1) Gillespie; k10; alpha0; nA0; nB0; kA0; kB0;
k30; beta0; n0; k0." << endl;
        exit(0);
    }

// Input arguments
//Number of cells in the population.
int N = atoi(args[1]);
// Frequency of environment switching.
double nu = atof(args[2]);
// Selection strength.
int sT = atoi(args[3]);
// Mutation rate (x individual x generation).
double u = atof(args[4]);
// Mutation factor.
double M = atof(args[5]);
// Algorithm: 0 = Deterministic , 1=Gillespie
bool alg = atoi(args[6]);
// Initial value for the parameters
double k10 = atof(args[7]);
double alpha0 = atof(args[8]);
double kA0 = atof(args[9]);
double kB0 = atof(args[10]);
double nA0 = atof(args[11]);
double nB0 = atof(args[12]);
double k30 = atof(args[13]);
double beta0 = atof(args[14]);
double k0 = atof(args[15]);
double n0 = atof(args[16]);

for(int iS=0; iS<numRep; iS++)
    {
// OUTPUT FILES
char myFileName[255];
// Average: Generation , Environment , Genotype
sprintf(myFileName," AGS_N%d_nu %1.4f_s%d_u %1.3f_M %1.2f_a

```

A. CÓDIGOS

```

    %d_k1 %1.3f_alpha %1.3f_kA %1.3f_kB %1.3f_nA %1.3f_nB
    %1.3f_k3 %1.3f_beta %1.3f_k %1.3f_n %1.3f_%d. dat" ,
    N,nu,sT,u,M,alg ,k10 ,alpha0 ,kA0,kB0,nA0,nB0 ,
    k30 ,beta0 ,k0 ,n0 ,iS );
    ofstream AGS(myFileName, ios::out );
    AGS << "Generation" << '\t' << "Environment" << '\t';
    AGS << "k1_T" << '\t' << "alpha_T" << '\t' << "kA_T"
    << '\t' << "kB_T" << '\t' << "nA_T" << '\t' << "nB_T"
    << '\t';
    AGS<< "k3_T" << '\t' << "beta_T" << '\t' << "k_T"
    << '\t' << "n_T" << '\t';
    AGS << "A_T" << '\t' << "B_T" << '\t' << "w_T" << endl;
    // Lineage information x cycle: Generation ,
    // Parental Lineage, #Mutations, Average Fitness, Genotype
    // in L and Genotype in H
    sprintf(myFileName,"LxC_N%d_nu%1.4f_s%d_u%1.3f_M%1.2f_a
    %d_k1 %1.3f_alpha %1.3f_kA %1.3f_kB %1.3f_nA %1.3f_nB %1.3f_k3
    %1.3f_beta %1.3f_k %1.3f_n %1.3f_%d. dat" ,
    N,nu,sT,u,M,alg ,k10 ,alpha0 ,kA0,kB0,nA0,nB0 ,k30 ,beta0 ,k0 ,
    n0 ,iS );
    ofstream LxC(myFileName, ios::out );
    LxC << "Generation" << '\t'; //Removed Bistable Flag
    LxC << "LTag" << '\t' << "LMut" << '\t' << "LFit" << '\t';
    LxC << "k1_H" << '\t' << "alpha_H" << '\t' << "kA_H"
    << '\t' << "kB_H" << '\t' << "nA_H" << '\t' << "nB_H"
    << '\t';
    LxC<< "k3_H" << '\t' << "beta_H" << '\t' << "k_H" << '\t'
    << "n_H" << '\t';
    LxC << "k1_L" << '\t' << "alpha_L" << '\t' << "kA_L"
    << '\t' << "kB_L" << '\t' << "nA_L" << '\t' << "nB_L"
    << '\t';
    LxC<< "k3_L" << '\t' << "beta_L" << '\t' << "k_L" << '\t'
    << "n_L" << endl;

    // Individuals information per epoch (first 9 &
    // last generations): Generation, Environment,
    // Cell tag, Genotype, parent Tag
    sprintf(myFileName,"IxE_N%d_nu%1.4f_s%d_u%1.3f_M%1.2f_a
    %d_k1 %1.3f_alpha %1.3f_kA %1.3f_kB %1.3f_nA %1.3f_nB %1.3f_k3
    %1.3f_beta %1.3f_k %1.3f_n %1.3f_%d. dat" ,N,nu,sT,u,M,alg ,k10 ,
    alpha0 ,kA0,kB0,nA0,nB0 ,k30 ,beta0 ,k0 ,n0 ,iS );
    ofstream IxE(myFileName, ios::out ); //Removed SS, added B
    IxE << "Generation" << '\t' << "Environment" << '\t'

```

```

<< "Cell_tag" << '\t';
IxE << "k1" << '\t' << "alpha" << '\t' << "kA" << '\t'
<< "kB" << '\t' << "nA" << '\t' << "nB" << '\t';
IxE<< "k3" << '\t' << "beta" << '\t' << "k" << '\t'
<< "n" << '\t';
IxE << "A" << '\t'<< "B" << '\t' << "w" << '\t';
IxE << "Parent_tag" << endl;

long seed = seeds[iS]; // Initialize random seed.
int E = 1 - E0; // Initialize environment.

// Create the cell population:
GeneNetwork myPop[N];
PopAverage myPopT; //PopAverage Total

// Assign user-defined-parameters values:
myPop[0].ic(A0,B0,k10, alpha0, nA0, nB0, kA0, kB0, k2, k30, beta0,
n0, k0, k4);
myPop[0].LTag = 0;
for (int i=1; i<N; i++)
{
    myPop[i] = myPop[0];
    myPop[i].LTag = i;
}
// Simulate generations:
for (int iG=1; iG<=GMAX; iG++)
{
// Update environmental state each epoch:
    if (iG%int(1/nu)==1)
    {
        E = (E+1)%2;
    }
myPopT.reset();
// Loop over the population:
for (int i = 0; i<N; i++)
{
// Gillespie Algorithm:
if (alg)
{
    while (myPop[i].time < TMAX)
    {
        myPop[i].updateTime(&seed);
        myPop[i].updateSystem(&seed);
    }
}
}
}

```

```

    }
}
else //Deterministic Algorithm
{
// initial conditions
state_type x = {myPop[i].A , myPop[i].B};
boost::numeric::odeint::integrate_adaptive(ctrl_rkck54(),
myPop[i] , x , 0.0 , tmax_deterministic , 0.001);
myPop[i].A=x[0];
myPop[i].B=x[1];
}
// Calculate the fitness of cell [i]:
myPop[i].calculateFitness(E, BOPT, V);
myPop[i].LFit += (myPop[i].w*(nu/2));
// TOTAL
myPopT.k1 += myPop[i].k1;
myPopT.alpha += myPop[i].alpha;
myPopT.kA += myPop[i].kA;
myPopT.kB += myPop[i].kB;
myPopT.nA += myPop[i].nA;
myPopT.nB += myPop[i].nB;
myPopT.k3 += myPop[i].k3;
myPopT.beta += myPop[i].beta;
myPopT.k += myPop[i].k;
myPopT.n += myPop[i].n;
myPopT.A += myPop[i].A;
myPopT.B += myPop[i].B;
myPopT.w += myPop[i].w;
myPopT.f++; //Fraction , in this case = N

// PRINTING – Individuals information per epoch
//(first 9 & last generations):
if((iG>(GMAX-(Ep/nu))) && (iG%iint(1/nu)<10))
{
IxE << iG << '\t' << E << '\t' << i << '\t';
IxE << myPop[i].k1 << '\t' << myPop[i].alpha << '\t' <<
myPop[i].kA << '\t' << myPop[i].kB << '\t' <<
myPop[i].nA << '\t' << myPop[i].nB << '\t';
IxE << myPop[i].k3 << '\t' << myPop[i].beta << '\t' <<
myPop[i].k << '\t' << myPop[i].n << '\t';
IxE << myPop[i].A << '\t' << myPop[i].B << '\t'
<< myPop[i].w << '\t';
IxE << myPop[i].parent << endl;

```

```

}
}

// PRINTING – Average genotype & subpopulations:
myPopT.normalize(N);
AGS << iG << '\t' << E << '\t';
AGS << myPopT.k1 << '\t' << myPopT.alpha << '\t' <<
myPopT.kA << '\t' << myPopT.kB << '\t' << myPopT.nA << '\t'
<< myPopT.nB << '\t';
AGS << myPopT.k3 << '\t' << myPopT.beta << '\t' << myPopT.k
<< '\t' << myPopT.n << '\t';
AGS << myPopT.A << '\t' << myPopT.B << '\t' << myPopT.w
<< endl;

// Each epoch, record genotype per lineage:
if (iG%int(1/nu)==0)
{
    for (int i=0; i<N; i++)
    {
        myPop[i].k1E[E] = myPop[i].k1;
        myPop[i].alphaE[E] = myPop[i].alpha;
        myPop[i].kAE[E] = myPop[i].kA;
        myPop[i].kBE[E] = myPop[i].kB;
        myPop[i].nAE[E] = myPop[i].nA;
        myPop[i].nBE[E] = myPop[i].nB;
        myPop[i].k3E[E] = myPop[i].k3;
        myPop[i].betaE[E] = myPop[i].beta;
        myPop[i].kE[E] = myPop[i].k;
        myPop[i].nE[E] = myPop[i].n;
    }
}

// Each cycle:
if (iG%int(2/nu)==0)
{
    //double AS[5] = {0,0,0,0,0};
    for (int i=0; i<N; i++)
    {
        if (iG>(GMAX-(Cp*2/nu)))
        {
            // PRINTING – Lineage information x cycle:
            LxC << iG << '\t';
            LxC << myPop[i].LTag << '\t' << myPop[i].LMut << '\t'
            << myPop[i].LFit << '\t';

```

```

    LxC << myPop[i].k1E[0] << '\t' << myPop[i].alphaE[0]
    << '\t' << myPop[i].kAE[0] << '\t' << myPop[i].kBE[0]
    << '\t' << myPop[i].nAE[0] << '\t' << myPop[i].nBE[0]
    << '\t';
    LxC << myPop[i].k3E[0] << '\t' << myPop[i].betaE[0]
    << '\t' << myPop[i].kE[0] << '\t' << myPop[i].nE[0]
    << '\t';
    LxC << myPop[i].k1E[1] << '\t' << myPop[i].alphaE[1]
    << '\t' << myPop[i].kAE[1] << '\t' << myPop[i].kBE[1]
    << '\t' << myPop[i].nAE[1] << '\t' << myPop[i].nBE[1]
    << '\t';
    LxC << myPop[i].k3E[1] << '\t' << myPop[i].betaE[1]
    << '\t' << myPop[i].kE[1] << '\t' << myPop[i].nE[1]
    << endl;
    }
    // Reset lineage's statistics:
    myPop[i].LMut = 0;
    myPop[i].LFit = 0;
    myPop[i].LTag = i;
    }
}
// Apply selection and evolve parameters:
selectionTournament(myPop, N, sT, u, M, &seed, PAR_E);
}
// Close files
    AGS.close();
    LxC.close();
    IxE.close();
}
return 0;
}

```

A.3. GeneNetwork_OscillatoryModel.h (Selección sobre B)

Lenguaje de programación: C++

```

// Evolutionary dynamics of gene network with potential for
// oscillations in fluctuating environments
// EVOLUTIONARY MODEL
//
// Created by Mariana Gomez-Schiavon
// October 2015

```

```
// Adapted for oscillatory model by Sophia Orozco
// October 2021

// Libraries :
#include <stdio.h>
#include <cmath>
#include <stdlib.h>
#include <time.h>
#include <iostream>
#include <algorithm>
#include <vector>
#include "ran1.h"

using namespace std;

class parE
{
public:
bool k1;
bool alpha;
bool nA;
bool nB;
bool kA;
bool kB;
bool k2;
bool k3;
bool beta;
bool n;
bool k;
bool k4;

parE()
{
k1 = 0;
alpha = 0;
nA = 0;
nB = 0;
kA = 0;
kB = 0;
k2 = 0;
k3 = 0;
beta = 0;
```


A. CÓDIGOS

```
        n = 0;
        k = 0;
        k4 = 0;
    }
};

class PopAverage
{
public:
    // Evolving parameters
    double k1, alpha, nA, nB, kA, kB, k3, beta, n, k;
    double A;        // Protein number
    double B;        // Protein number
    double w;        // Fitness
    double f;        // Fraction

    void reset()
    {
        k1 = 0;
        alpha = 0;
        nA = 0;
        nB = 0;
        kA = 0;
        kB = 0;
        k3 = 0;
        beta = 0;
        n = 0;
        k = 0;
        A = 0;
        B = 0;
        w = 0;
        f = 0;
    }

    void normalize(int N)
    {
        k1 = k1/f;
        alpha = alpha/f;
        nA = nA/f;
        nB = nB/f;
        kA = kA/f;
        kB = kB/f;
        k3 = k3/f;
    }
};
```

```

        beta = beta/f;
        n = n/f;
        k = k/f;
        A = A/f;
        B = B/f;
        w = w/f;
        f = f/N;
    }
};

class GeneNetwork
{
public:
    // Genotype:
    double k1, alpha, nA, nB, kA, kB, k3, beta, n, k;
    // Evolving parameters
    double k2, k4; //Fixed parameters
    // Phenotype:
    double A; //Protein number
    double B; //Protein number
    // Evolutionary metrics:
    double w; //Fitness
    int parent; //Parent cell number
    // Flag for new mutant individuals
    //(0 – Parental genotype; 1 – Mutant)
    bool mut;
    // Lineage metrics:
    int LTag; //Tag for parental lineage
    int LMut; //Number of accumulated mutas per cycle
    double LFit; //Average fit of the lineage in cycle
    double k1E[2]; //Genotype k1 at the end of each epoch
    double alphaE[2]; //Genotype alpha at the end of e epoch
    double nAE[2]; //Genotype nA at the end of each epoch
    double nBE[2]; //Genotype nB at the end of each epoch
    double kAE[2]; //Genotype kA at the end of each epoch
    double kBE[2]; //Genotype kB at the end of each epoch
    double k3E[2]; //Genotype k3 at the end of each epoch
    double betaE[2]; //Genotype beta at the end of e epoch
    double nE[2]; //Genotype n at the end of each epoch
    double kE[2]; //Genotype k at the end of each epoch

    // Gene dynamics variables:
    // Reaction propensities & total reaction propensity

```

A. CÓDIGOS

```
double r[4], rT;
double time;          // Running time
//2 Variable System (Deterministic)
typedef boost::array< double , 2 > state_type;

    GeneNetwork()
    {
        w = 0;
        parent = 0;
        mut = 0;
        LMut = 0;
        LFit = 0;
        time = 0;
    }

// Assign initial genotype & phenotype:
void ic(double A0, double B0, double k10, double alpha0,
double nA0, double nB0, double kA0, double kB0, double k20,
double k30, double beta0, double n0, double k0, double k40)
{
A = A0;          // Protein number
B = B0;          // Protein number
k1 = k10;       // Maximum synthesis rate
alpha = alpha0; // Basal synthesis activity
nA = nA0;       // Hill coefficient
nB = nB0;       // Hill coefficient
kA = kA0;       // Affinity constant
kB = kB0;       // Affinity constant
k2 = k20;       // Protein degradation constant
k3 = k30;       // Maximum synthesis rate
beta = beta0;   // Basal synthesis activity
n = n0;         // Hill coefficient
k = k0;         // Affinity constant
k4 = k40;       // Protein degradation constant
}

//Stochastic solution
// Update reaction propensities:
void updateR()
{
// A synthesis reaction
r[0] = k1*(alpha*(pow(kB,nB)/(pow(kB,nB)+pow(B,nB)))+
(1-alpha*(pow(kB,nB)/(pow(kB,nB)+pow(B,nB))))*(pow(A,nA)
```

```

/(pow(A,nA)+(pow(kA,nA)/pow(kB,nB))
*(pow(B,nB)+pow(kB,nB)))));
r[1] = k2*A; // A degradation reaction
// B synthesis reaction
r[2] = k3*(beta+((1-beta)*(pow(A,n)/(pow(k,n)+pow(A,n)))));
// B degradation reaction
r[3] = k4*B;
// Total propensity
rT = r[0]+r[1]+r[2]+r[3];
    }

// Perform the Gillespie algorithm:
void updateTime(long *seed)
{
    // Random number between 0 and 1
    double rand1 = (double)ran1(seed);
    // Time between reactions
    double re_time;
        updateR(); // Update reaction propensities
        // Time for the next reaction
        re_time = -log(1-rand1)/rT;
        time += re_time; // Actualize global time
    }
void updateSystem(long *seed)
{
    // Random number between 0 and 1
    double rand2 = (double)ran1(seed);
    if (rand2 < r[0]/rT)
    {
        A++; // A Synthesis reaction
    }
    else if (rand2 <= (r[0]+r[1])/rT)
    {
        A--; // A Protein degradation
    }
    else if (rand2 <= (r[0]+r[1]+r[2])/rT)
    {
        B++; // B Synthesis reaction
    }
    else if (rand2 <= (r[0]+r[1]+r[2]+r[3])/rT)
    {
        B--; // B Protein degradation
    }
}

```

```

        else
        {
            cerr << "Runtime Error: rT not real."
            << endl;
            exit(0);
        }
    }

//Deterministic
void operator() ( const state_type &x ,
state_type &dxdt , double t )
{
dxdt[0] = k1*(alpha*(pow(kB,nB)/(pow(kB,nB)
+pow(x[1],nB)))+(1-alpha*(pow(kB,nB)/(pow(kB,nB)
+pow(x[1],nB))))*(pow(x[0],nA)/(pow(x[0],nA)+
(pow(kA,nA)/pow(kB,nB))*
(pow(x[1],nB)+pow(kB,nB)))))-k2*x[0];
dxdt[1] = k3*(beta+((1-beta)*(pow(x[0],n)/
(pow(k,n)+pow(x[0],n)))))-k4*x[1];
}

// Evaluate cell's fitness:
void calculateFitness(int myE, double Bopt[2], double v)
{
    // Lorentzian fitness function.
    w = 1/(1+((B-Bopt[myE])*(B-Bopt[myE])/(v*Bopt[myE])));
}

// Random mutation in the genotype:
void mutate(long *seed, double M, parE PAR_E)
{
    if(PAR_E.k1)
    {
        k1 *= pow(M,1-(2*ran1(seed)));
        if(k1>1000) //Upper limit
        {
            k1=1000;
        }
        if(k1<0.01) //Lower limit
        {
            k1=0.01;
        }
    }
}

```

```
if (PAR.E.alpha)
{
    alpha *= pow(M,1 - (2*ran1(seed)));
    if (alpha >= 1)
    {
        alpha = 0.999;
    }
    if (alpha < 0.001)
    {
        alpha = 0.001;
    }
}
if (PAR.E.nA)
{
    nA *= pow(M,1 - (2*ran1(seed)));
    if (nA > 16)
    {
        nA = 16;
    }
    if (nA < 0.01)
    {
        nA = 0.01;
    }
}
if (PAR.E.nB)
{
    nB *= pow(M,1 - (2*ran1(seed)));
    if (nB > 16)
    {
        nB = 16;
    }
    if (nB < 0.01)
    {
        nB = 0.01;
    }
}
if (PAR.E.kA)
{
    kA *= pow(M,1 - (2*ran1(seed)));
    if (kA > 100)
    {
        kA = 100;
    }
}
```

```
        if (kA<0.01)
        {
            kA=0.01;
        }
    }
    if (PAR.E.kB)
    {
        kB *= pow(M,1 - (2*ran1(seed)));
        if (kB>100)
        {
            kB = 100;
        }
        if (kB<0.01)
        {
            kB=0.01;
        }
    }
    if (PAR.E.k3)
    {
        k3 *= pow(M,1 - (2*ran1(seed)));
        if (k3>1000)
        {
            k3=1000;
        }
        if (k3<0.01)
        {
            k3=0.01;
        }
    }
    if (PAR.E.beta)
    {
        beta *= pow(M,1-(2*ran1(seed)));
        if (beta>=1)
        {
            beta = 0.999;
        }
        if (beta<0.001)
        {
            beta=0.001;
        }
    }
    if (PAR.E.n)
    {
```

```

        n *= pow(M,1-(2*ran1(seed)));
        if (n>16)
        {
            n=16;
        }
        if (n<0.01)
        {
            n=0.01;
        }
    }
    if (PAR.E.k)
    {
        k *= pow(M,1-(2*ran1(seed)));
        if (k>100)
        {
            k = 100;
        }
        if (k<0.01)
        {
            k=0.01;
        }
    }
    mut = 1;
    LMut++;
}

};

// TOURNAMENT SELECTION
void selectionTournament(GeneNetwork *myPop, int N, int t,
double u, double M, long *seed, parE PAR.E)
{
    // The new population to be defined
    GeneNetwork newPop[N];
    // Fill in the next generation:
    int iN = 0;
    int rand1;
    while (iN < N)
    {
        // Stochastically choose the tournament cells and
        //replicate the cell with higher fitness in the group:
        int indexT[t];
        indexT[0] = int(N*(double)ran1(seed));

```


A. CÓDIGOS

```
int maxT = indexT[0];
int j = 1;
while (j < t)
{
rand1 = int(N*(double)ran1(seed));
for (int i=0; i < j; i++)
{
    if (rand1 == indexT[i])
    {
        rand1 = -1;
        break;
    }
}
if (rand1 > 0)
{
    indexT[j] = rand1;
    if(myPop[maxT].w < myPop[indexT[j]].w)
    {
        maxT = indexT[j];
    }
    j++;
}
}
//The new cell is equal to the old cell selected.
newPop[iN] = myPop[maxT];
newPop[iN].parent = maxT;    // Parent cell number
iN++;
}

// Refill the population matrix & allow random mutation:
for (int i = 0; i < N; i++)
{
myPop[i] = newPop[i];
myPop[i].time = 0;           // Reset the time to zero.
myPop[i].mut = 0;           // Flag for parental genotype.
if ((double)ran1(seed) <= u)
{
    // Random number between 0 and 1
    double mutationalscenario = (double)ran1(seed);
    if(mutationalscenario <=0.25) //Protein B
    {
        PAR.E.k1 = 0;
        PAR.E.alpha =0;
    }
}
}
}
```

```
    PAR.E.nA = 0;
    PAR.E.nB = 1;
    PAR.E.kA = 0;
    PAR.E.kB = 1;
    PAR.E.beta =0;
    PAR.E.k3 = 0;
    PAR.E.n = 0;
    PAR.E.k = 0;
}
else if(mutationalscenario <=0.5) //Promoter A
{
    PAR.E.k1 = 1;
    PAR.E.alpha =1;
    PAR.E.nA = 1;
    PAR.E.nB = 1;
    PAR.E.kA = 1;
    PAR.E.kB = 1;
    PAR.E.beta =0;
    PAR.E.k3 = 0;
    PAR.E.n = 0;
    PAR.E.k = 0;
}
else if(mutationalscenario <=0.75) //Protein A
{
    PAR.E.k1 = 0;
    PAR.E.alpha =0;
    PAR.E.nA = 1;
    PAR.E.nB = 0;
    PAR.E.kA = 1;
    PAR.E.kB = 0;
    PAR.E.beta =0;
    PAR.E.k3 = 0;
    PAR.E.n = 1;
    PAR.E.k = 1;
}
else //Promoter B
{
    PAR.E.k1 = 0;
    PAR.E.alpha =0;
    PAR.E.nA = 0;
    PAR.E.nB = 0;
    PAR.E.kA = 0;
    PAR.E.kB = 0;
```

A. CÓDIGOS

```
        PAR.E.beta =1;
        PAR.E.k3 = 1;
        PAR.E.n = 1;
        PAR.E.k = 1;
    }
    myPop[ i ].mutate( seed , M, PAR.E);
}
}
```

Bibliografía

- [1] Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461. [1](#), [4](#), [13](#)
- [2] Barkai, N. and Leibler, S. (2000). Circadian clocks limited by noise. *Nature*, 403(6767):267–268. [14](#)
- [3] Bell-Pedersen, D., Cassone, V. M., Earnest, D. J., Golden, S. S., Hardin, P. E., Thomas, T. L., and Zoran, M. J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics*, 6(7):544–556. [17](#)
- [4] Bernhardt, J. R., O’Connor, M. I., Sunday, J. M., and Gonzalez, A. (2020). Life in fluctuating environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1814):20190454. [16](#), [17](#)
- [5] Çağatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J., and Süel, G. M. (2009). Architecture-Dependent Noise Discriminates Functionally Analogous Differentiation Circuits. *Cell*, 139(3):512–522. [48](#)
- [6] Cai, L., Friedman, N., and Xie, X. S. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358–362. [12](#)
- [7] Eldar, A. and Elowitz, M. B. (2010). Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–173. [12](#), [13](#), [48](#)
- [8] Elowitz, M. B. and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338. [14](#)
- [9] Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. (2002). Stochastic Gene Expression in a Single Cell. *Science*, 297(5584):1183–1186. [12](#)
- [10] Fasani, R. A. and Savageau, M. A. (2010). Automated construction and analysis of the design space for biochemical systems. *Bioinformatics*, 26(20):2601–2609. [24](#)
- [11] Ferrell, J. E., Tsai, T. Y.-C., and Yang, Q. (2011). Modeling the Cell Cycle: Why Do Certain Circuits Oscillate? *Cell*, 144(6):874–885. [10](#), [11](#)

BIBLIOGRAFÍA

- [12] Foster, P. L. (2005). Stress responses and genetic variation in bacteria. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 569(1-2):3–11. [16](#)
- [13] Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361. [12](#), [27](#)
- [14] Golding, I., Paulsson, J., Zawilski, S. M., and Cox, E. C. (2005). Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell*, 123(6):1025–1036. [12](#)
- [15] Gomez-Schiavon, M. (2016). Stochastic Dynamics and Epigenetic Regulation of Gene Expression: from Stimulus Response to Evolutionary Adaptation. page 252. [6](#), [16](#), [33](#)
- [16] Gómez-Schiavon, M. and Buchler, N. E. (2019). Epigenetic switching as a strategy for quick adaptation while attenuating biochemical noise. *PLOS Computational Biology*, 15(10):e1007364. [1](#), [8](#), [18](#), [27](#), [32](#), [33](#), [59](#), [63](#)
- [17] Gonze, D. and Abou-Jaoudé, W. (2013). The Goodwin Model: Behind the Hill Function. *PLoS ONE*, 8(8):e69573. [7](#), [10](#)
- [18] Gonze, D. and Goldbeter, A. (2006). Circadian rhythms and molecular noise. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(2):026110. [15](#), [47](#)
- [19] Gonze, D., Halloy, J., and Goldbeter, A. (2004). Emergence of coherent oscillations in stochastic models for circadian rhythms. *Physica A: Statistical Mechanics and its Applications*, 342(1-2):221–233. [15](#), [47](#)
- [20] Goodwin, B. C. (1965). Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation*, 3:425–437. [10](#)
- [21] Kim, J. K. (2016). Protein sequestration versus Hill-type repression in circadian clock models. *IET Systems Biology*, 10(4):125–135. [11](#)
- [22] Ko, C. H., Yamada, Y. R., Welsh, D. K., Buhr, E. D., Liu, A. C., Zhang, E. E., Ralph, M. R., Kay, S. A., Forger, D. B., and Takahashi, J. S. (2010). Emergence of Noise-Induced Oscillations in the Central Circadian Pacemaker. *PLoS Biology*, 8(10):e1000513. [16](#)
- [23] Li, Q. and Lang, X. (2008). Internal Noise-Sustained Circadian Rhythms in a Drosophila Model. *Biophysical Journal*, 94(6):1983–1994. [15](#), [47](#), [63](#), [64](#)
- [24] Li, Z., Liu, S., and Yang, Q. (2017). Incoherent Inputs Enhance the Robustness of Biological Oscillators. *Cell Systems*, 5(1):72–81.e4. [14](#)
- [25] Losick, R. and Desplan, C. (2008). Stochasticity and Cell Fate. *Science*, 320(5872):65–68. [13](#), [17](#)
- [26] Luria, S. E. and Delbrück, M. (1943). MUTATIONS OF BACTERIA FROM VIRUS SENSITIVITY TO VIRUS RESISTANCE. *Genetics*, 28(6):491–511. [16](#)

- [27] Meyers, L. A. and Bull, J. J. (2002). Fighting change with change: adaptive variation in an uncertain world. *Trends in Ecology & Evolution*, 17(12):551–557. [16](#), [17](#)
- [28] Novák, B. and Tyson, J. J. (2008). Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991. [1](#), [10](#)
- [29] Raj, A. and van Oudenaarden, A. (2008). Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell*, 135(2):216–226. [11](#), [12](#)
- [30] Salathé, M., Van Cleve, J., and Feldman, M. W. (2009). Evolution of Stochastic Switching Rates in Asymmetric Fitness Landscapes. *Genetics*, 182(4):1159–1164. [18](#)
- [31] Savageau, M. A. and Fasani, R. A. (2009). Qualitatively distinct phenotypes in the design space of biochemical systems. *FEBS Letters*, 583(24):3914–3922. [24](#)
- [32] Strogatz, S. H. (1994). *Nonlinear dynamics and Chaos: with applications to physics, biology, chemistry, and engineering*. Studies in nonlinearity. Addison-Wesley Pub, Reading, Mass. [10](#), [23](#)
- [33] Süel, G. M., Garcia-Ojalvo, J., Liberman, L. M., and Elowitz, M. B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. *Nature*, 440(7083):545–550. [48](#), [63](#)
- [34] Thattai, M. and van Oudenaarden, A. (2004). Stochastic Gene Expression in Fluctuating Environments. *Genetics*, 167(1):523–530. [18](#)
- [35] Turner, T., Schnell, S., and Burrage, K. (2004). Stochastic approaches for modeling in vivo reactions. *Computational Biology and Chemistry*, 28(3):165–178. [12](#)
- [36] Tyson, J. J. and Novák, B. (2010). Functional Motifs in Biochemical Reaction Networks. *Annual Review of Physical Chemistry*, 61(1):219–240. [1](#), [9](#)
- [37] Valderrama-Gómez, M. Á., Lomnitz, J. G., Fasani, R. A., and Savageau, M. A. (2020). Mechanistic Modeling of Biochemical Systems without A Priori Parameter Values Using the Design Space Toolbox v.3.0. *iScience*, 23(6):101200. [26](#), [39](#)
- [38] Vilar, J. M. G., Kueh, H. Y., Barkai, N., and Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *Proceedings of the National Academy of Sciences*, 99(9):5988–5992. [15](#), [47](#)
- [39] Watson, J. D., editor (2014). *Molecular biology of the gene*. Pearson, Boston, seventh edition edition. [3](#), [4](#)