



Modelo matemático para las estrategias de replicación de plásmidos mediante juegos estocásticos

Daniel Eduardo Fajardo Fajardo

Universidad de Los Andes
Facultad de Ciencias, Departamento de Física
Bogotá D.C., Colombia
Noviembre, 2017



Modelo matemático para las estrategias de replicación de plásmidos mediante juegos estocásticos

Daniel Eduardo Fajardo Fajardo

Proyecto de Grado
Física

Asesor:
Juan Manuel Pedraza, PhD

Universidad de Los Andes
Facultad de Ciencias, Departamento de Física
Bogotá D.C., Colombia
Noviembre, 2017

"Anybody who looks at living organisms knows perfectly well that they can produce other organisms like themselves. This is their normal function, they wouldn't exist if they didn't do this, and it's not plausible that this is the reason why they abound in the world. In other words, living organisms are very complicated aggregations of elementary parts, and by any reasonable theory of probability or thermodynamics highly improbable. That they should occur in the world at all is a miracle of the first magnitude; the only thing which removes, or mitigates, this miracle is that they reproduce themselves. Therefore, if by any peculiar accident there should ever be one of them, from there on the rules of probability do not apply, and there will be many of them, at least if the milieu is reasonable. But a reasonable milieu is already a thermodynamically much less improbable thing. So, the operations of probability somehow leave a loophole at this point, and it is by the process of self-reproduction that they are pierced. "

John von Neumann, Theory Of Self Reproducing Automata

Agradecimientos

Quiero agradecerle a mi familia por estar siempre en los momentos difíciles.
Tampoco olvido a Diegol por su ayuda incondicional durante este proceso tan
complejo.

Resumen

En este documento se realiza un análisis numérico de la relación que hay entre la carga metabólica de los plásmidos que se relaciona con el crecimiento de la bacteria, el nivel de rigidez del control de copias y el número esperado de copias con la razón final de las poblaciones de plásmidos mediante una simulación de Montecarlo. Por otro lado, se obtiene una expresión fenomenológica que describe la relación entre los parámetros anteriores. Esto se consigue a través de la descripción de un proceso de Morán para simular la reproducción de los plásmidos en la bacteria. También se explora una primera aproximación por redes bayesianas para un primer paso en el modelo teórico. [1, 2]

Índice general

1. Introducción	8
2. Fundamentos Conceptuales	10
2.1. Fundamento económico	10
2.2. Fundamento biológico	14
2.2.1. Proceso de Moran	14
2.2.2. Plásmidos	14
2.3. Fundamento computacional	24
2.3.1. Método de Montecarlo	24
3. Metodología	25
4. Resultados y Análisis	27
4.0.1. Redes Bayesianas	27
4.0.2. Simulación a través de las generaciones	28
4.0.3. Acerca del fit	29
4.0.4. Barrido de k	30
4.0.5. Barrido de h	30
4.0.6. Barrido de b	31
4.0.7. Fórmula fenomenológica	32
5. Conclusiones	33
6. Apéndice	34

Índice de figuras

2.1. Probabilidad de replicación según ecuación de Hill para control laxo	22
2.2. Probabilidad de replicación según ecuación de Hill para control estricto	23
4.1. Gráfica logarítmica de las poblaciones en función de las generaciones	29
4.2. Gráfica de razones de los logaritmos de las poblaciones en función de las generaciones	29
4.3. Gráfica de razones de los logaritmos de las poblaciones en función de la razón de los k	30
4.4. Gráfica de razones de los logaritmos de las poblaciones en función de h	31
4.5. Gráfica de razones de los logaritmos de las poblaciones en función de b	32

Capítulo 1

Introducción

La teoría de juegos es una rama de la matemática aplicada que estudia el conflicto y la cooperación de distintos actores racionales, lo cual implica que tratan de aumentar su ganancia.[3]. Esto puede abordar sistemas biológicos tanto como sistemas sociales compuestos por humanos.

La facultad de Economía está trabajando interdisciplinariamente con el departamento de Física para estudiar el comportamiento de escenarios en los cuales se quieren establecer analogías directas entre el juego del bien común y el juego de los plásmidos en la bacteria en lo que a la teoría de juegos respecta. En economía se dispone de un juego que se describe por el bien común, y los jugadores enfrentan la decisión de cooperar o ser egoístas en el grupo. El juego consiste en que cada jugador posee un capital inicial y puede decidir si conservarlo en un fondo privado o invertirlo en el fondo común: El dinero en el fondo común se multiplica por un factor mayor a uno y se reparte en partes iguales a todos los jugadores, de tal manera que el juego tendría un equilibrio de Nash en un comportamiento totalmente egoísta, mientras el óptimo social estaría en la cooperación absoluta. Lo interesante sucede cuando el factor del fondo común se hace mayor para grupos más cooperativos lo cual hace que haya competencia a dos niveles: individual(jugadores) y grupal (grupo de jugadores). En biología hay unas relaciones que establecen cuándo es evolutivamente estable la cooperación.[1] Lo cual es bastante similar al problema que queremos analizar en biología, donde distinguimos dos organismos: plásmidos (jugadores) y bacterias (grupo de jugadores). Los plásmidos viven al interior de las bacterias y son trozos de ADN extracromosomal que se replican con la maquinaria de la célula.

El presente trabajo es un modelo que simula el comportamiento a nivel biológico de plásmidos en las bacterias, y es una primera aproximación para lograr un modelo analítico de la situación con el fin de encontrar resultados sin necesidad de una simulación. A futuro se quiere verificar la validez del modelo mediante trabajo experimental: una vez establecida la validez del modelo se podrá comparar con los resultados obtenidos en la Facultad de Economía y encontrar equivalencias entre los distintos parámetros en los diferentes entornos. La simulación utilizará como base el proceso de Moran para simular la reproducción de los plásmidos, esto mediante el método de Montecarlo para tener en cuenta la estocasticidad de la reproducción. Se variarán los parámetros de la simulación para mediante regresiones encontrar una fórmula fenomenológica de los resultados obtenidos en función de los parámetros. En esta etapa del proyecto total se establecen algunas analogías entre el modelo biológico y el juego del bien común que se hace en la Facultad de Economía, aunque vale la pena decir que no son las más refinadas. En el planteamiento económico del problema se abordó de manera supremamente breve una primera aproximación, por redes bayesianas, un método popular para resolver juegos de información incompleta.

Capítulo 2

Fundamentos Conceptuales

En este documento se trata un tema interdisciplinario que requiere una fundamentación teórica en varios ámbitos, como el económico, cuando se hace referencia a la teoría de juegos, en el biológico y en las herramientas teóricas de la simulación utilizada.

2.1. Fundamento económico

Basado en [3]

La teoría de juegos es una disciplina formal de la matemática aplicada que estudia y modela la interacción entre racionales en situaciones de competencia y cooperación. En teoría de juegos es muy popular el dilema del prisionero. Este fue propuesto por Flood y Dresher, pero solo fue formalizado por Tucker. Consiste en la siguiente situación: durante un robo un guardia es asesinado y los dos ladrones responsables son capturados. A los ladrones se les ofrece un trato: si un jugador delata al otro, pero el otro no delata, tiene la opción de salir libre inmediatamente pero el otro debe pagar una sentencia de 10 años. Si ambos se delatan mutuamente deben pagar 7 años. Pero si ambos se quedan callados pagan 1 año de cárcel. Estos posibles resultados se pueden sintetizar en una matriz de pagos 2.1 que se ve a continuación. Según la teoría de juegos el resultado lógico es que ambos jugadores hagan su mejor jugada suponiendo que los demás jugadores hacen lo mejor para sus intereses. En este caso el equilibrio de Nash es delatar al otro. Este equilibrio de Nash difiere del óptimo social. Lo interesante es ver cuando en un juego entran a importar los mecanismos de cooperación porque la solución del juego converge

		Jugador Y	
		<i>Cooperar</i>	<i>Delatar</i>
Jugador X	<i>Cooperar</i>	(1, 1)	(10, 0)
	<i>Delatar</i>	(0, 10)	(7, 7)

Cuadro 2.1: Matriz de pagos

al óptimo social y por eso es importante mencionar este juego. Para poner un ejemplo un poco más aterrizado a lo que se quiere hacer en esta monografía es bueno analizar la tragedia de los comunes. Consiste en un juego de varios jugadores donde está la opción de ser egoísta y abusar de los recursos compartidos para maximizar las ganancias individuales, pero si todos los jugadores juegan de esa manera los recursos comunes se agotan. Entonces para evitar una catástrofe lo mejor es cooperar a cierta escala para evitar un colapso. Un ejemplo aterrizado a la vida real es el uso desmesurado que le hemos dado a los recursos naturales que nos ha llevado a una gran extinción de la fauna y flora. 1.3.1Juego del bien común

Basado en [4]

El trabajo de Isaac et al. (1994) busca responder una serie de preguntas relacionadas con el juego de los bienes comunes, relacionadas todas con la validez de los resultados experimentales en escenarios más realistas. En los experimentos, dicen los autores, normalmente participan pocos individuos, reunidos en un laboratorio, los cuales en un período aproximado de dos horas juegan el número de rondas que el experimentador encuentra conveniente (10 rondas, normalmente), y posteriormente se les recompensa monetariamente sus ganancias de puntos obtenidos durante el juego. Lo anterior implementación, si bien resulta práctica en materia logística, presenta serias limitaciones, como la imposibilidad de conocer si los resultados se mantienen o modifican a mayor número de participantes. Ahora bien, cabe precisar que el juego de los bienes comunes reseñado por los autores en el trabajo, parte del diseño experimental de Marwell y Ames (1979,1980) y que también manejan Miller y Andreoni (1993), e Isaac y Walter (1998), los cuales entienden el juego como un mecanismo de distribución de la dotación (Z) del individuo i entre dos fondos, uno privado y uno público, los cuales presentan rendimientos diferentes. Por un lado, el privado rinde siempre p_i , mientras que el fondo público, el cual depende no sólo de la inversión del jugador i sino también del resto de jugadores, rentará $[G(m_i + m_j)]/n$, donde m_i es la inversión del jugador i en el fondo público, m_j es la inversión del resto de jugadores en este fondo, y n es el número de participantes. (Isaac et al., pág. 3). Así, lo que todos

invierten, se multiplica por un factor G y se distribuye igualmente entre todos los participantes, independientemente de si estos aportaron al fondo común o no.

Dado el diseño planteado, controlando que $\pi_i = Z < m_i = Z$ y que el retorno marginal per cápita de la cuenta pública, o $MPCR > 1$, se puede concluir que la situación óptima en el sentido de Pareto sería para el jugador i invertir todo en el fondo público, mientras que el Equilibrio de Nash se encuentra en invertirlo todo en el fondo privado. Entonces, teóricamente se podría esperar que los jugadores invirtieran todo en el bien privado a la par que esperaran beneficiarse de las contribuciones de otros jugadores en la cuenta pública, haciendo que éstos actuaran como “polizones” (en inglés, free rider). Sin embargo, los resultados parecen demostrar comportamientos distintos. Para todos los experimentos reseñados por los autores, siempre existía una cifra promedio positiva de aportes a las cuentas públicas, pero siendo esta distante del óptimo de Pareto. Teniendo en cuenta las variables controladas por Isaac y Walter (IW) (1988) para intentar explicar este fenómeno, se encuentran el tamaño de la población y el MPCR. Partiendo de esta base, los autores del acá reseñado artículo idean un nuevo planteamiento metodológico que busca entender si en grandes grupos, mucho más grandes que los planteados por IW (grupos de 10 vs grupos de 4 personas), con una variación sobre MPCR igual a la señalada por estos, y bajo otros esquemas de recompensas, se pueden llegar a los mismos resultados que señalan IW sobre la influencia positiva del MPCR sobre los aportes y la no relevancia del tamaño poblacional o si, por el contrario, se corroborará la premisa teórica ampliamente aceptada, que predice el aumento de polizones a mayor número de individuos interactuando. Así, Isaac et al. proponen la formación de varios grupos de 4, 10, 40 y 100 personas, siendo los de 4 y 10 considerados como “grupos pequeños” y los restantes dos “grupos grandes”. Los participantes juegan en condiciones de $MPCR = 0,30$ o $MPCR = 0,75$, durante varios días, y con incentivos a la participación no monetarios, sino académicos. Ya que los sujetos son estudiantes, se les recompensa con nota extra, ahorrando así valiosos recursos. De este experimento, se obtiene que, en primer lugar, los grupos pequeños efectivamente se comportan como lo señalado por IW, respondiendo estos positivamente a los aumentos en MPCR, y sin un cambio significativo al pasar de un grupo de 4 personas a uno de 10 personas. Así, los individuos, sin importar el tamaño del grupo en el que se encuentren, deciden cooperar en promedio más y en la misma proporción en aquellos grupos que juegan bajo condiciones de retornos marginales de 0,75 sobre aquellos grupos que juegan bajo MPCR del 0,30. Sin embargo, sí encuentran una diferencia significativa entre los considerados grupos grandes y los pequeños, siendo más eficientes los grupos grandes, lo que desvirtúa la premisa de a mayor tamaño, mayor número

de polizontes. Sin embargo, lo anterior sólo es cierto bajo $MPCR = 0,30$, ya que a $MPCR = 0,75$ las diferencias son poco significativas.

Otra conclusión relevante hallada por los autores es que, en los grupos grandes, un cambio en el $MPCR$, al menos entre los niveles de 0,30 a 0,75 no afecta la proporción promedio invertida en el bien público. Ahora bien, los datos obtenidos también encuentran la existencia de una inversión promedio positiva en la cuenta pública, lo que demuestra una desviación del equilibrio de Nash, el cual teóricamente los jugadores deberían buscar. Isaac et al. (1994) proponen varias explicaciones a este fenómeno. Dentro de las más relevantes, se encuentran el aprendizaje sobre el equilibrio, las preferencias representadas de forma incorrecta, y las expectativas sobre las siguientes rondas. Sobre el segundo punto, menciona que el juego no representa las preferencias individuales por la “justicia” o el sentimiento moral de aportar a la cuenta pública, por lo que no capta las reales funciones de utilidad individuales. Al respecto, Ledgard (1993) en su selección de grupos, realiza un proceso de criba, donde separa según el nivel de satisfacción social, obteniendo resultados que tienden a mostrar una relación negativa entre la cantidad de polizontes y la satisfacción social.

Sobre el último punto, las expectativas a futuro, los autores incluyen la expectativa de éxito del individuo al señalar su intención de aportar a la cuenta pública en la ronda actual (i). La probabilidad de éxito se entiende si jugador i logra, en la ronda inmediatamente siguiente, que $p_i + m_i + \sum m_{j*} = Z$, donde m_{j*} representa el número de tokens que se deben aportar como mínimo para lograr que la suma de $=Z$. Así, este m_{j*} , para cada jugador j , sería igual a $m_{j*}/(n-1)$. Dado lo anterior, se deduce que, a mayor número de personas, mayor la expectativa de influenciar en los otros una decisión de aportar, ya que el mínimo de tokens que deben aportar cada individuo para que i logre Z es menor. Al compararlo con los resultados, se aprecia que efectivamente esta podría ser una explicación acertada. Finalmente, y como consideración importante, los autores mencionan que, dado que se influencia a futuro, al acercarse al final del juego los aportes deberían disminuir. Sin embargo, los resultados muestran que inclusive en la última ronda existen contribuciones a la cuenta pública, lo que muestra la complejidad de las decisiones tomadas por las personas.

Notese que aunque el juego fue definido inicialmente para dinero y grupos de personas, el conflicto que modela puede verse reflejado en otras situaciones. Por ejemplo, los jugadores pueden ser compañías y el bien común infraestructura, o como veremos mas adelante, los jugadores pueden ser otros seres vivos y el bien común recursos bioquímicos en el ambiente.

2.2. Fundamento biológico

2.2.1. Proceso de Moran

Un proceso de Moran, nombrado así en honor a Patrick Moran que lo propuso por primera vez en 1958, es un proceso estocástico para simular poblaciones finitas y es útil para este trabajo, pues nos ayuda con el problema de dos alelos en una población, o mejor dicho, una mutación, que es lo que se quiere resolver. Este algoritmo solo tiene en cuenta el fitness, que es que tan adecuado es el individuo para nacer y sobrevivir, para el nacimiento, es decir que al momento de nacer que escoge aleatoriamente que organismo nace pero es una distribución donde hay distintos pesos y el peso es lo que llamamos fitness y la muerte es completamente aleatoria en el sentido en el que no hay pesos. Algoritmos más complejos tienen en cuenta el fitness tanto en el nacimiento como en la muerte. Aquí es importante el tamaño de la población pues afecta la diferencia entre el pay-off relativo y el pay-off total.

2.2.2. Plásmidos

Basado en [5][6]

Los plásmidos son moléculas extracromosomales de ADN lineal o circular, presentes principalmente en Bacteria y Archaea. Estos plásmidos mantienen un número de copias característico en su hospedero, y pueden llegar a constituir un porcentaje importante de su ADN. Más aún, estos plásmidos son capaces de incorporar o transferir genes desde o hacia su hospedero a través de eventos de recombinación o conjugación, favoreciendo así los intercambios genéticos en poblaciones bacterianas.

El interés científico por los plásmidos deriva de la gran cantidad de información genética que pueden almacenar, así como de su potencial impacto en comunidades microbianas, y su potencial para ser empleados como vectores de clonación. Su investigación se ve facilitada por tres factores: pueden ser fácilmente aislados y manipulados, tienen una organización genética sencilla, y no son esenciales para la supervivencia de su hospedero (por lo que se en principio se pueden introducir y retirar de éste sin afectar al último). No obstante, dichos elementos también resultan de interés en su carácter de elementos egoístas, que emplean la maquina-

ria de replicación del hospedero para replicarse y repararse. En ese sentido, genes que mejoren la capacidad de supervivencia de las bacterias (como aquellos que otorgan resistencia a antibióticos) pueden verse como herramientas del plásmidos para asegurar su multiplicación.

Mecanismos de replicación de plásmidos

Basado en [5][6] Existen tres formas de replicación en los plásmidos circulares: (i) replicación tipo Theta, (ii) replicación por desplazamiento de cadena, y (iii) replicación por círculos rodantes. Por su parte, los plásmidos lineales que se han encontrado tanto en bacterias Gram-positivas como negativas, emplean una de dos estrategias: (i) aquellos plásmidos que poseen horquillas en sus extremos se replican a través de intermediarios concatémicos, mientras que (ii) los que tienen proteínas unidas a su extremo 5' se replican a través de cebadores cuya síntesis es mediada por proteínas.

El material genético de los plásmidos puede dividirse en dos partes: • Una región esencial, que contiene todos los genes y secuencias nucleotídicas requeridas para su replicación y regulación, la cual se compone de: (i) un origen de replicación u ori, i.e., el segmento específico del plásmido donde inicia la replicación, (ii) aquellos genes asociados al inicio de la replicación; y (iii) aquellos genes codificantes de proteínas encargadas de la regulación de la replicación. • Una región no esencial, compuesta por el resto del plásmido. Esta contiene diversos genes, encargados, por ejemplo, de la transferencia y dispersión del plásmido en la población, o de la resistencia a antibióticos o metales pesados.

Replicación Tipo Theta Bajo el mecanismo Theta, la replicación ADN inicia con la síntesis de un primer de ARN, seguida de su extensión covalente. Así como en los organismos celulares, la replicación es continua en la hebra líder y discontinua en la rezagada. La replicación puede ser unidireccional o bidireccional, e iniciar en uno o varios orígenes. En la mayoría de los casos, el inicio de la replicación precisa una proteína iniciadora, codificada por el plásmido y denominada Rep, además de la ADN polimerasa III del hospedero.

Secuencia ori y Proteínas Rep La secuencia ori en la replicación tipo Theta puede definirse de tres maneras: (i) es la mínima región con acción en cis que permite la replicación autónoma; (ii) es la región donde las hebras de ADN se separan para iniciar la replicación; o (iii) corresponde a la(s) base(s) donde inicia la replicación de la hebra líder.

Generalmente, las secuencias ori están conformadas por tres regiones: (i) una zona de unión de las proteínas iniciadoras codificadas por el plásmido (si son necesarias); (ii) una región adyacente repetitiva y con abundante contenido AT, donde se da la separación de las hebras; y (iii) uno o varios sitios (cajas dnaA) para la unión de la proteína iniciadora DnaA del hospedero. Se debe tener en cuenta que en la secuencia ori puede haber sitios de unión de otros factores requeridos para el comienzo de la replicación.

Las regiones donde las proteínas Rep se unen, usualmente están formadas por grupos de secuencias repetitivas –adyacentes o no– denominadas iterones. Estos pueden hallarse también fuera de la región ori, y pueden llegar a jugar un papel importante en la regulación de la replicación dado que una proteína Rep se une a un solo iterón en ori. Las proteínas Rep pueden actuar como monómeros o dímeros dependiendo del caso, y en algunas veces se requieren ambas conformaciones para iniciar la replicación.

Estas proteínas pueden participar en la regulación de la replicación. Las interacciones proteína-proteína de las proteínas Rep son mediadas por motivos tipo LZ (cremalleras de leucina), mientras que las interacciones proteína-ADN son mediadas por motivos HTH (hélice-giro-hélice).

Iniciación y Terminación de la Replicación

La iniciación de la replicación involucra el ensamblaje de la maquinaria completa de replicación, la cual incluye principalmente las enzimas ADN polimerasa III, helicasa DnaB, y una primasa. Las proteínas Rep, que se unen al plásmido en ori, están a cargo de la separación de la doble cadena en la zona rica en AT, lo cual posibilita la síntesis del primer de ARN. Nótese, sin embargo, que algunos plásmidos no emplean proteínas Rep durante la iniciación.

Después del comienzo de la replicación de la hebra líder, el proceso continúa hasta su terminación con la participación ocasional de otras proteínas del hospedero. El sitio específico de terminación se conoce como ter. La terminación suele implicar interferencia con la actividad de la helicasa, enzima encargada de mantener la separación de las hebras. Proteínas ancladas a ter son las que previenen que el proceso de replicación continúe. Para la hebra rezagada, la terminación parece estar a cargo de RNAs o proteínas que interfieren con la maquinaria de replicación.

Cuando el proceso se completa, los plásmidos producidos pueden estar entrelazados entre sí. Las topoisomerasas tipo I o II del hospedero son las enzimas que

se encargan de su separación.

Replicación por Desplazamiento de Cadena Contrario al mecanismo Theta, el desplazamiento de cadena involucra la síntesis simultánea de ambas hebras en direcciones opuestas, desplazando a la otra hebra. Adicionalmente, se requieren tres proteínas de iniciación codificadas por el plásmido –RepA , RepB, y RepC– para la iniciación de la replicación.

Secuencia ori y Proteínas Rep La secuencia ori en la replicación por desplazamiento de cadena se define como la región mínima que permite la replicación bidireccional cuando se suministra el complejo de replicación, i.e., las proteínas Rep. De forma similar a la zona ori del mecanismo Theta, la región mínima incluye tres zonas: (i) tres iterones idénticos de 20-bp; (ii) un segmento con una región rica en AT y otra en GC; y (iii) dos secuencias palindrómicas, una en cada hebra, denominadas ssiA y ssiB. Las primeras dos zonas se requieren para la separación de la doble cadena, en tanto que las regiones ssi están involucradas en la síntesis del primer de ARN.

En conclusión, las tres proteínas Rep, RepA, RepB, y RepC, son, en ese orden, una helicasa, una primasa, y un iniciador. Sus roles en la replicación se explican más adelante.

Iniciación y Terminación de la Replicación La iniciación no requiere la acción de proteínas iniciadoras codificadas por el hospedero, pero sí precisa el uso de la ADN polimerasa III, y otras proteínas del hospedero con capacidad de unión al ADN.

El proceso inicia cuando RepC se une a los iterones en ori. Acto seguido, RepA se une al ADN en un sitio aledaño a la región rica en AT, y se desplaza a lo largo de la cadena, de forma tal que produce la separación de las dos hebras. Esto permite a RepB sintetizar el primer de ARN en uno de los sitios ssi.

El proceso lleva a la generación de una molécula de ADN de doble cadena y otra de cadena sencilla. La cadena complementaria de la segunda es sintetizada a partir de la generación de un primer de ARN en su zona ssi.

Replicación por Círculo Rodante En este tipo de replicación el proceso es asimétrico y unidireccional: la hebra líder se replica independientemente de la rezagada. Este mecanismo es generalmente usado por plásmidos pequeños (menos de 10kb). El proceso inicia con el cleavage del ADN del plásmido en una región específica conocida como dso. Usando el 3'-OH generado como primer, la replica-

ción inicia probablemente a través de proteínas codificadas por el hospedero. La elongación continúa desplazando la otra cadena hasta que se alcanza de nuevo el sitio dso, momento en el cual ocurre la terminación.

Secuencia ori y Proteínas Rep En la replicación por círculo rodante la región ori que se replica debe contener tres elementos: (i) la ya mencionada zona dso, (ii) un gen rep el cual codifica la proteína Rep, y (iii) los genes de los elementos involucrados en el control de la replicación. El plásmido debe contener uno o dos zonas sso las cuales son precisadas para la síntesis de la hebra complementaria al ADN de cadena sencilla producida tras la replicación.

El dso está compuesto por dos regiones: una secuencia conservada nick capaz de formar horquillas y donde se cliva el ADN; y otra más variable denominada bind, donde se une la proteína Rep. Ambas están compuestas de secuencias repetitivas.

Las proteínas Rep de estos plásmidos tiene actividad ADN transferasa, es decir, tienen la habilidad de clivar y ligar ADN.

Iniciación y Terminación de la Replicación La iniciación de la replicación de la hebra líder involucra la formación de una horquilla en la región nick de dso. La proteína Rep, actuando como dímero, se une a bind y cliva el ADN en algún punto del bucle de la horquilla.

Pese a que se sabe poco del complejo proteico a cargo de la elongación, la ADN polimerasa III parece liderar el proceso. Como se mencionó previamente, la terminación ocurre al momento de volver a pasar por (y replicar) la región dso. Allí, la proteína Rep, la cual sigue unida al ADN, cataliza la ligación del sitio dso original con el otro extremo de la misma molécula, produciendo un intermediario de ADN de cadena sencilla. La proteína Rep también se encarga de ligar la hebra que acaba de sintetizarse, generando una molécula de ADN de doble cadena.

La conversión del subproducto de cadena sencilla a cadena doble se inicia en la región sso, la cual también parece tener la capacidad de formar horquillas. No obstante, los detalles del proceso aún no se conocen en su totalidad. Se sospecha que es la maquinaria de replicación del hospedero la que se encarga de la conversión, requiriendo la síntesis de primers de ARN.

Mecanismos de Regulación de la Replicación

Basado en [5][6] Existen tres mecanismos de regulación de la replicación:

i. ARNs anti-sentido: este ARN puede unirse directamente a la región del ARN original que codifica una secuencia, o se une al ADN, o puede unirse en otra zona y generar un cambio conformacional que inactive el ARN. Este mecanismo se observa generalmente en bacterias.

Los ARNs que se regulan de esta forma son primers usados en la replicación o ARNs mensajeros de proteínas asociadas al inicio de la replicación (incluyendo proteínas Rep).

ii. Una combinación de ARNs anti-sentido y proteínas: el principio de los ARNs anti-sentido no cambia, mientras que las proteínas actúan como represores transcripcionales, a veces regulándose a sí mismas. ¿Ventaja de tener proteínas?

iii. Sitios en el ADN afines a proteínas reguladoras: corresponde a la regulación por iterones, escenario en el que las proteínas Rep se autorregulan a sí mismas, de tal forma que logran una concentración óptima. ¿Cómo se compara/contrasta con los otros dos mecanismos? Se sabe que las proteínas Rep tienen la capacidad de unirse a más de un iterón a la vez. Si la concentración de Rep es alta, entonces estas se unirán a iterones de distintas moléculas produciendo un impedimento estérico que evita la replicación.

Además de estos tres mecanismos de regulación, el proceso de metilación también puede estar involucrado en el control de la replicación de plásmidos, bien inhibiendo o estimulando el proceso.

Control de numero de copias

Basado en [5][6]

La expresión de los genes o plásmidos determinan la diferenciación de las células y es un proceso altamente regulado. Es muy importante que los genes sean expresados tanto en el tipo correcto de células como a un nivel adecuado cuando la célula responde a señales internas y externas. La regulación de la expresión de genes puede ocurrir de diferentes maneras: (i) durante la transcripción, (ii) durante el procesamiento del mARN, o (iii) durante la traducción al nivel de la estabilidad de la proteína. Sin embargo, se cree que la regulación ocurre principalmente en la

transcripción.

La maquinaria de transcripción de procariotas consiste en dos componentes regulatorios complementarios: el cis-acting y el trans-acting. Los elementos de cis-acting son secuencias de ADN que crean secciones accesibles para que los factores trans se enlacen en la transcripción inicial, mientras que los elementos trans-acting son factores de transcripción que reconocen y se ligan a secuencias específicas en los elementos cis-acting para iniciar, acelerar o suprimir la transcripción. El número de copias que se producen de un gen está determinado por la velocidad con la que los elementos trans-acting se adhieren a los elementos cis-acting, ya que este proceso afecta la iniciación, la aceleración o la supresión del proceso de transcripción del gen. Es importante tener en cuenta que los sitios de enlace trans-cis se llaman promotores y hacen parte de los elementos cis-acting.

Una característica esencial de los plásmidos es la habilidad de replicarse como elementos genéticos autónomos de forma controlada dentro del hospedero. Son genes que se autorreplican y que se encuentran generalmente en los procariotas.

Conectando los aspectos moleculares y ecológicos se puede mostrar que la replicación de plásmidos se enfrenta a una selección a dos niveles. Por un lado, están los plásmidos que se replican demasiado en el interior de la célula, haciendo que haya mayor posibilidad de dejar descendencia en células hijas pero disminuyendo su velocidad de crecimiento dada la alta carga metabólica que se presenta. Por otro lado, están los plásmidos que se replican en el interior de las células sin tener en cuenta el interior de las otras células.

Por ende, los plásmidos se enfrentan a tres sistemas de control: el proceso de auto-replicación de los plásmidos, la selección intracelular debido a la compatibilidad de los plásmidos, y la selección intercelular en el crecimiento. La interacción de estos tres factores puede llegar a acelerar o suprimir el proceso de replicación.

Los plásmidos definen su propia capacidad de replicación al controlar la iniciación de la replicación y el número en la célula hospedero. Si hay demasiadas copias replicadas, la célula crecerá lentamente debido a la sobre carga metabólica. Los plásmidos evitan llegar a esto.

El control del inicio de la replicación se debe a activadores e inhibidores. Los activadores son capaces de interactuar con cis o trans dependiendo del plásmido, con el origen de la replicación es atraer ADN polimerasa e iniciar la replicación, mientras que los inhibidores actúan solo en los trans e interrumpen la línea de producción del activador. Esto se puede expresar como una ecuación cinética de

una sola razón.

Si dos tipos de plásmidos se replican dentro del mismo hospedero, la célula solo se fijará en el número total de plásmidos y no en el número de cada tipo. En este caso, la probabilidad de que un plásmido pueda replicarse más rápido que el otro es más alta, y cuando la división celular se lleva a cabo, un plásmido de rápida replicación se transferirá más que el otro.

Los plásmidos compiten con las células para la replicación puesto que deben equilibrar dos situaciones opuestas: deben minimizar el número de copias con el fin de evitar una mayor carga metabólica, y al tiempo deben evitar desaparecer en el caso que su número no sea suficientemente alto. Es importante recordar que la carga metabólica depende de condiciones externas, expresión de genes y número de copias.

Los plásmidos dentro de un huésped controlan su tasa de reproducción segregando una sustancia inhibidora (proteína o ARN) que desacelera el crecimiento de esta población, pero la expresión de este inhibidor es un proceso estocástico. La probabilidad (P) en función del número de plásmidos (n) puede modelarse como una función de Hill, donde los parámetros " h " y " k " pueden variar dependiendo del gen (Ver ecuación 2.4). El k viene relacionado con la tasa de rapidez con la que se sintetiza el inhibidor, también con su vida media y con la tasa de replicación media[2]. Hay mutaciones que disminuyen la susceptibilidad del plásmido al inhibidor, aumentan el número de copias y por tanto favorecen a los portadores de este gen como individuos. Sin embargo, si el huésped se llena de este tipo de plásmidos, el costo metabólico hace que la bacteria hospedera crezca más lentamente y la población dentro del huésped se vea perjudicada. Si graficamos la probabilidad de reproducción de los plásmidos en función de la concentración, obtendríamos un resultado similar al de las curvas de la gráfica 1, donde la función amarilla corresponde a un número de copias menor y la función azul concierne a menor susceptibilidad al inhibidor y, por tanto, un número promedio de copias del plásmido mayor. Este comportamiento o estrategia en teoría de juegos permite un mayor beneficio a escala individual, lo que implica un equilibrio de Nash. Pero la cuestión es que si todos juegan esa estrategia, el grupo fracasa. Ahí es donde entran mecanismos de cooperación como los descritos por Nowak (2006)[7]. Si un plásmido decide cooperar, es decir, no desinhibirse, paga un costo a nivel individual y es que su probabilidad de reproducirse disminuye, pero el costo que paga es un sacrificio que debe balancearse con el beneficio para la población. Este costo puede variar dependiendo de las formas de las funciones de respuesta al inhibidor. Por ejemplo, una función más empinada corresponde a un control

más estricto del número de copias (en la gráfica 2 vemos el equivalente al caso 1 para un control más estricto del número de copias). Reducir la variabilidad en el número de copias tiene beneficios para el plásmido, pero resulta en una competencia más fuerte entre mutantes de distinto número promedio de copias. Esto hace que el costo de cooperar dependa directamente de la función de respuesta al inhibidor. La diferencia entre las figuras 2.1 y 2.2 ilustra este efecto: en el primer caso, para un número de plásmidos igual a 5, los plásmidos naranjas con control tienen menor probabilidad de reproducirse que los azules; mientras que en el caso 2 para esta concentración definitivamente son los azules los que se reproducen y los naranjas no tienen ninguna posibilidad. Esto se refleja directamente en los costos evolutivos de cambiar de curva. Un objetivo a largo plazo del laboratorio es corroborar los resultados de Paulsson (2002) y Nowak (2006) [6] [7] para describir esta situación y que establecen bajo qué condiciones es evolutivamente estable un comportamiento altruista (Ecuación 2). Este resultado es notable porque los detalles del comportamiento se reducen a unos pocos parámetros efectivos. Sin embargo, la medición de dichos parámetros es difícil, y en particular el cálculo del término del costo requiere un modelaje del juego evolutivo ya que tiene que tener en cuenta el efecto estocástico de la reproducción. Lo interesante en este proyecto es encontrar esa función de costo basado en las funciones de Hill (Ecuación 1) que modelan la probabilidad de reproducción, teniendo en cuenta un juego de varias rondas hasta que se establezca una estrategia dominante.

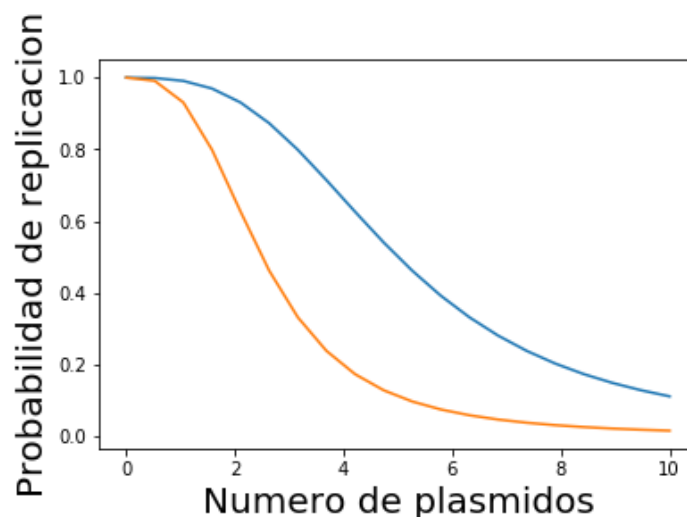


Figura 2.1: Probabilidad de replicación según ecuación de Hill para control laxo

La ecuación 2.4 se puede demostrar a partir de las reacciones químicas que suceden donde P es el plásmido, R el represor que crean los plásmidos y Q el

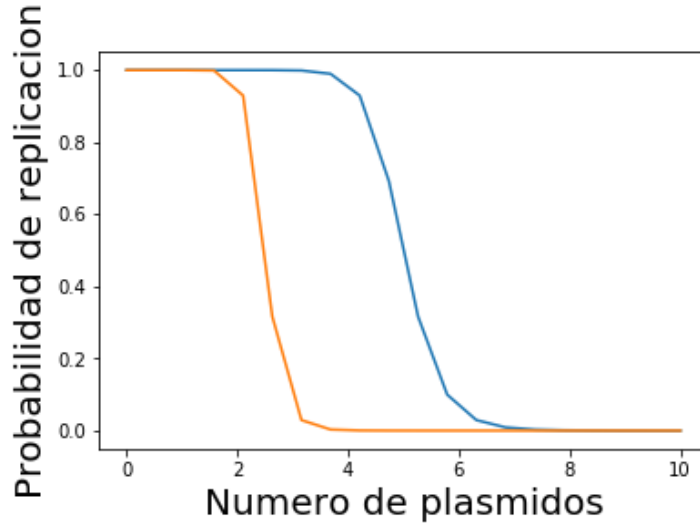


Figura 2.2: Probabilidad de replicación según ecuación de Hill para control estricto

motor de la bacteria que replica un plásmido:[6]. La dinámica de la proteína inhibidora no es tenida en cuenta directamente ya que la síntesis y adherencia de este al plásmido sucede en un tiempo de orden mucho menor a la de la dinámica del plásmido como tal. Entonces es válido solo preocuparse por la dinámica de reproducción del plásmido dada por la función de Hill. Esta probabilidad está dada para un tiempo discreto. Y toda la simulación se hace con esta discretización del tiempo.

$$P- > P + R$$

(2.1)

$$P + R->PR$$

(2.2)

$$Q + P \rightarrow 2P + Q \quad (2.3)$$

$$Probabilidad = \frac{1}{1 + (k * n)^h} \quad (2.4)$$

Ecuación 2.4. Probabilidad de reproducción (P) en función de número de plásmidos(n), parámetro de sensibilidad a inhibidor (k), exponente de Hill que determina si es un control laxo o estricto del número de plásmidos(h).

$$T = Tasadecrecimiento = \frac{a}{b + n} \quad (2.5)$$

Ecuación 2.5 muestra la tasa crecimiento de la bacteria, que es inversamente proporcional al número de plásmidos "n", donde a y b son constantes que se asignan y corresponden biológicamente para darle mayor o menor peso al número de plásmidos en la tasa de crecimiento.

2.3. Fundamento computacional

2.3.1. Método de Montecarlo

Este método es utilizado para simular situaciones no deterministas cuya solución exacta es muy costosa de evaluar con exactitud, consiste en asignar un número aleatorio a la variable aleatoria siguiendo determinada distribución y con ello resolver una ecuación determinista y realizar muchas iteraciones para promediar y así obtener una solución aproximada. Antes se utilizaba para juegos de azar, ahora se usa para casi todo Pues su error decae como $1/\sqrt{N}$, donde N es el número de intentos.

Capítulo 3

Metodología

Para la simulación de una bacteria se tienen plásmidos de dos tipos, una población mutada que empieza siendo una minoría y una población dominante inicialmente. La población mutada tiene un k menor que le permite reproducirse con mayor facilidad de acuerdo a la ecuación 2.4, pues para determinar si un plásmido se reproduce se elige un número aleatorio y si el número al azar es menor que la probabilidad dada por la ecuación 2.4, el plásmido se duplica. Esto sucede en una unidad de tiempo y ese tiempo se suma en el tiempo de vida de esa bacteria. Una bacteria tiene tres atributos: el número de plásmidos tipo uno, el número de plásmidos tipo dos y el tiempo de vida. Una bacteria se divide cuando alcanza cierto tamaño, y la razón de crecimiento depende del número de plásmidos de acuerdo a la ecuación 2.5, dado que existe un costo metabólico de tener plásmidos dentro de una bacteria. Al dividirse la bacteria hereda su tiempo actual de vida a sus hijas y los plásmidos en su interior se reparten siguiendo una distribución binomial. Esto se hace con un número aleatorio y una función cumulativa de la binomial respectiva con $p=0.5$. Las bacterias que se reproducen son aquellas con menor tiempo de vida. Esto se repite en un número igual a $2^{generaciones}$. A este tipo de simulación se le llama simulación de procesos de Moran [8] En el apéndice se incluye el código en Python utilizado. A continuación explico los métodos del código: Se define una función "binomial" que recibe dos enteros como entrada y calcula la binomial con $p=0.5$. Se define una función "cumbin" que calcula las cumulativas de una binomial en un array para cierto n . Se define una función que recibe como entrada el número de plásmidos 1, el número de plásmidos 2, sus respectivos "k", el exponente de Hill "hz", "az" "b". Y simula un proceso de crecimiento de una bacteria, donde devuelve el tiempo en el que la bacteria alcanza el tamaño en el cual se divide y el número de cada plásmido. Luego se hace el método que realiza la repartición binomial. Finalmente se hace el método

realiza este procesos un gran número de veces pero solo reproduce a las bacterias con menor tiempo de reproducción.

Capítulo 4

Resultados y Análisis

Se hizo una simulación poniendo a competir a dos tipos de plásmidos, un plásmido dominante con una población de diez y un ejemplar mutado con un parámetro k menor, lo cual le da una facilidad para reproducirse superior a la del plásmido dominante, se escogió diez por facilidad con los cálculos pues los cálculos funcionaron bien con ese número desde el principio.

4.0.1. Redes Bayesianas

A medida que pasa el tiempo, los plásmidos deciden si reproducirse o no basándose en la probabilidad de la ecuación 2.4 y la bacteria crece dada la razón de crecimiento de la ecuación 2.5. La probabilidad de que los plásmidos dentro de una bacteria se puede modelar teniendo en cuenta que la probabilidad de reproducción de un plásmido es la ecuación 4.1. Si se define un turno para la reproducción de cada tipo de plásmido se podría escribir como en la ecuación 4.2, donde los superíndices denotan la ronda. De esta manera basta con aplicar el método de Montecarlo con una sola función binomial, en vez de aplicar muchas Bernoulli. Esto sucede porque el plásmido tiene dos posibilidades en cada fracción de tiempo: reproducirse o no con una probabilidad dada por la ecuación 2.4. Si tomamos los eventos independientes de cada plásmido en un turno, se concluye que es una binomial, como se muestra más adelante. Pero a pesar de ello resulta más pesado computacionalmente hacer esto, pues se requiere hacer una acumulativa binomial y esto requiere más tiempo. Pero podría ser un comienzo para la solución analítica. Este enfoque se puede ver como una red de posibilidades, más exactamente una

Cuadro 4.1: Parámetros de la primera simulación

k1	k2	h	b
0.4	0.2	10	100

red bayesiana.

$$P(n_1, n_2) = \frac{1}{1 + (k * (n_1 + n_2))^h} \quad (4.1)$$

$$Probabilidad(n_1 \rightarrow n_1^0 | n_2^0) = \binom{n_1^0}{n_1^1 - n_1^0} (P(n_1^0, n_2^0))^{n_1^1 - n_1^0} * (1 - P(n_1^0, n_2^0))^{2n_1^0 - n_1^0} \quad (4.2)$$

De ahora en adelante se harán simulaciones con un número inicial de plásmidos tipo 1 de 10 y de plásmido tipo 2 de 1. El exponente de Hill utilizado es 10 porque se acerca a lo encontrado en la biología experimental[2] aunque el dato no es exacto porque difieren dependiendo del sistema es una muy buena aproximación y el k del mutado es la mitad del no mutado. Esto ya es una exageración necesaria para ver los resultados en un número computable de generaciones, donde el tipo tiene un parámetro k que le permite tener un mayor número de copias mientras menor sea. El parámetro a siempre es 100 veces menor que b , la razón a/b representa el tiempo que tarda una bacteria en dividirse, porque si b fuera mayor el crecimiento sería muy lento y la simulación demasiado pesada. Se simulan únicamente dos genotipos debido a que el tiempo que tarda un tipo de plásmidos en establecer una dominancia completa es muy corto comparado con el tiempo en el que aparacen nuevas mutaciones. El crecimiento se tomó como algo determinista porque es una muy buena aproximación a la realidad, y se tiene en cuenta la escasez de los recursos en la bacteria.

4.0.2. Simulación a través de las generaciones

El resultado sugiere que las bacterias portadoras de plásmidos quedan con una población homogénea en su interior, es decir, plásmidos de tipo 1 solamente, plásmidos de tipo 2 solamente, o vacías, lo que sugiere un crecimiento exponencial que se puede ver a través de las generaciones como se puede apreciar en la imagen 4.1. Los parámetros son tabulados en la tabla 4.3 Pero es interesante ver la razón entre las bacterias donde domina el tipo 1 y el tipo 2. Esto permite ver que con el paso de las generaciones es una cantidad que converge a un valor rápidamente dependiendo de los parámetros. El estudio de esta razón en términos de estos parámetros permite encontrar una relación interesante.

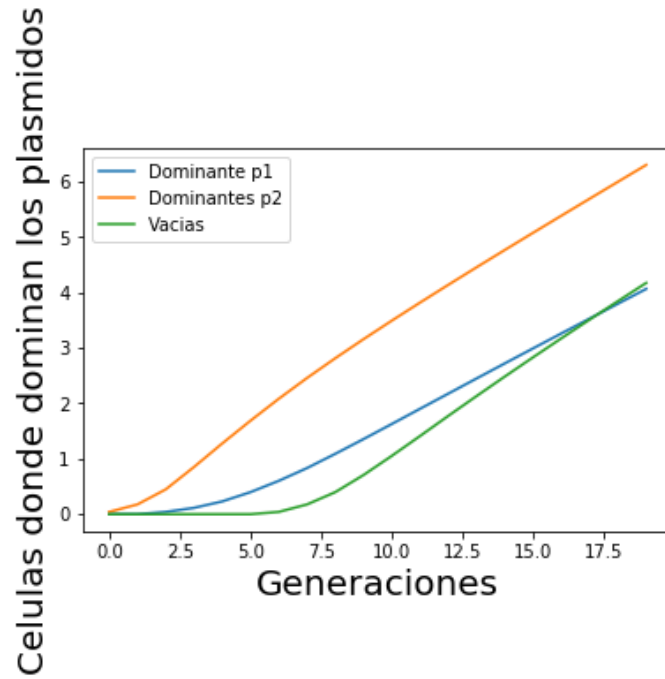


Figura 4.1: Gráfica logarítmica de las poblaciones en función de las generaciones

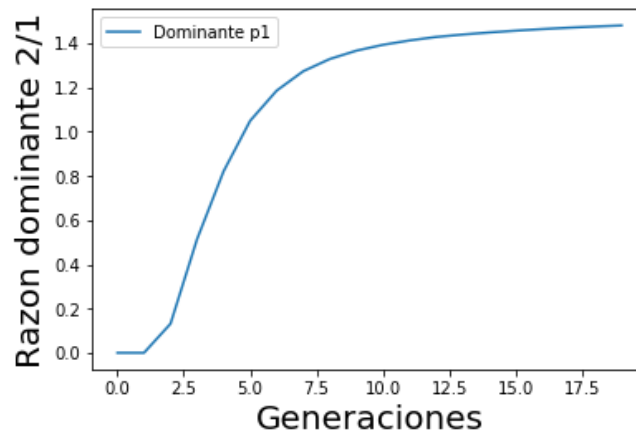


Figura 4.2: Gráfica de razones de los logaritmos de las poblaciones en función de las generaciones

4.0.3. Acerca del fit

Para encontrar el mejor fit se hizo una regresión lineal con el parámetro y con algunas potencias del parámetro. Aquel que maximizara el r era el elegido.

4.0.4. Barrido de k

En la figura 4.3 se puede ver que a medida que se aumenta la razón $\frac{k_2}{k_1}$, también se aumenta la razón de la población de plásmidos tipo 2 respecto al tipo 1. Este resultado tiene mucho sentido biológico pues a medida que se aumenta el número de copias permitido por la nueva población, esta establecerá su dominancia con mayor grado de eficacia. Así, se llega a la conclusión de que el exponente que se ajusta a ese comportamiento es 1, por ende es una relación lineal, es decir, $Razon \propto \frac{k_2}{k_1}$

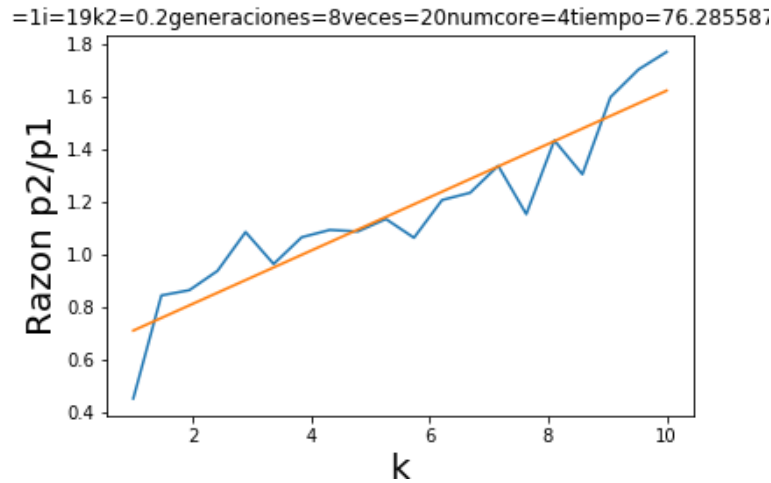


Figura 4.3: Gráfica de razones de los logaritmos de las poblaciones en función de la razón de los k

4.0.5. Barrido de h

En la figura 4.4 se ve que al aumentar el exponente de Hill "h" la razón de población 2 contra 1 disminuye, al hacer el análisis del mejor fit se encuentra que es -0.25 . $Razon \propto h^{-\frac{1}{4}}$ Este resultado va en contra de lo que indica la intuición, pues a medida que se aumenta el exponente de Hill resulta que la ventaja del plásmido con mayor número esperado de copias se hace más pronunciada y se esperaría que la razón aumente en lugar de disminuir. Los parámetros son tales que k_1 va desde 0.2 hasta 2 para ver que sucede desde que son iguales hasta 10 veces mayor, de nuevo exagerando lo que sucede cuando muta.

Cuadro 4.2: Parámetros de la primera simulación

k1	k2	h	b
0.2-2	0.2	10	100

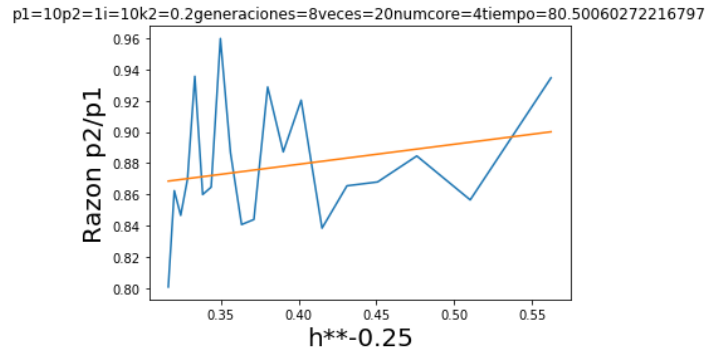


Figura 4.4: Gráfica de razones de los logaritmos de las poblaciones en función de h

4.0.6. Barrido de b

En la figura 4.5 se puede ver que a medida que aumentó el parámetro b , que implica una disminución en la importancia de la carga metabólica de los plásmidos en el crecimiento de la bacteria, se ve que la razón entre los tipo 2 y los tipo 1 aumenta. El mejor fit arrojó que el exponente era 0.5. Este resultado tiene sentido biológico pues se espera que, al castigar menos el comportamiento egoísta, este se haga más popular entre los plásmidos. Aún así, es un crecimiento lento pues a nivel individual el plásmido egoísta tiene ventaja a nivel individual, pero a nivel grupal, las bacterias con demasiados plásmidos crecen más lento y la combinación de ambas circunstancias dejan claro que no se verá tan pronunciado el crecimiento de la razón. La simulación se corre con unos valores de b desde 50 hasta 5000, para ver los extremos donde importa mucho en la razón de crecimiento el número de plásmidos a cuando importa muy poco.

Cuadro 4.3: Parámetros de la simulación barrido de b

k1	k2	h	b
0.4	0.2	10	50-5000

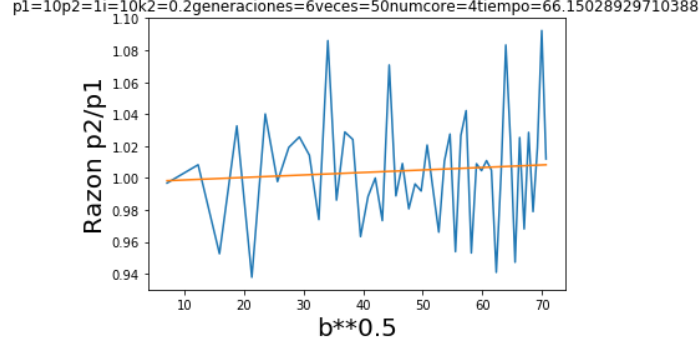


Figura 4.5: Gráfica de razones de los logaritmos de las poblaciones en función de b

4.0.7. Fórmula fenomenológica

Al unir los resultados de todos los barridos se concluye la siguiente fórmula fenomenológica

$$Razon2/1 \propto b^{\frac{1}{2}} * \frac{k1}{k2} * h^{\frac{-1}{4}} \quad (4.3)$$

La analogía con el juego de la Facultad de Economía es que la razón de $k2/k1$ es que tan egoístas son las personas que son egoístas en comparación con los altruistas. Pues el k es un coeficiente que viene representando cuál es el número esperado de copias. El coeficiente de Hill es qué tan laxo o estricto es el control, es decir, para un control estricto se tiene que castiga más a las personas altruistas en comparación a las egoístas que en el control estricto. Y el coeficiente b es que tanta importancia se le da a que en un grupo haya más cooperativas cuando se asigna el coeficiente por el cual se va a multiplicar el dinero del fondo común.

Capítulo 5

Conclusiones

- Mediante un análisis meticuloso de las mejores regresiones se propone que la formula fenomenológica es la siguiente:

$$Razon2/1 \propto b^{\frac{1}{2}} * \frac{k1}{k2} * h^{\frac{-1}{4}} \quad (5.1)$$

- Se encontró que al aumentar el parámetro h la razón disminuye, éste resultado era algo que no se esperaba pues es contrario a la intuición ya que se intuye que al aumentar lo estricto que es el control se cree que aumentará la dominancia del plásmido tipo 2 respecto al 1. Pero si se mira más a fondo lo que se hace es hacer más clara la linea que divide a las tipologías de los dos jugadores, entonces a nivel de la competencia grupal se establece na dominancia de los propiamente cooperativos
- La variación de los otros dos parámetros dio algo consistente con la intuición biológica, pues mostró una relación proporcional al aumentar el parámetro b y $\frac{k2}{k1}$
- En el juego del bien común la analogía con el parametro b indica que a medida que aumentamos el coeficiente por el cual se multiplica el fondo común, su raíz cuadrada es directamente proporcional con el indice en el cual aumentan los grupos que cooperan cada vez más y la razón de los k tiene que ver con que tanto se ven perjudicados los altruistas al aumentar el nivel de trampa de aquellos que no cooperan. Además, vemos que la estrategia evolutivamente estable con es ni la cooperación absoluta (óptimo social) ni el egoísmo absoluto (equilibrio de Nash) sino un punto intermedio.

Capítulo 6

Apéndice

In []:

```
In [3]: import matplotlib
#matplotlib.use('Agg')
import matplotlib.pyplot as plt
from math import factorial
from math import fabs
from math import floor
import random as random
import operator as op
import numpy as np
from math import log10
import time as time
import multiprocessing as mp
from joblib import Parallel, delayed
```

```
In [4]: def binomial(n,r):

    return ncr(n,r)/(2.0**n)
#Metodo que define una distribucion binomial con p=0.5

    def ncr(n, r):

        return factorial(n)/(factorial(r)*factorial(n-r))
#Metodo que devuelve n combinado k
```

```
In [5]: def cumbin(n):
    cum=0.0
    cumn1 = []
    for i in range(n):
        cum+=binomial(n,i)
        cumn1.append(cum)
    return cumn1
#Metodo que devuelve un array guardando la cumulativa de la binomial
```

```
In [6]: import random

def plasmidos(p1,p2,h,k1,k2,a,b,tiempo):
    #print (tiempo)

    def dRv(a,b,p):
        pfinal=floor(1/k1+1/k2+1)
        return a/((b+p))
    Rv=1.0
    def r1(plas1,plas2):
        p=1.0/(1.0+((plas1+plas2)*k1)**h)

        if random.random()<p:
            return True
        else: return False
```

```

def r2(plas1,plas2):
    p=1.0/(1.0+((plas1+plas2)*k2)**h)
    if random.random()<p:
        return True
    else: return False
pfinal=floor(1/k1+1/k2+1)
rondas=floor(10*(b+(pfinal))/a)
for i in range(rondas):
    if Rv>2:
        break
    if random.random()<0.5:
        for j in range(p1):
            if Rv>2:
                break
            if r1(p1,p2):
                p1=p1+1
            Rv=Rv+dRv(a,b,p1+p2)
            tiempo=tiempo+1

        for l in range(p2):
            if Rv>2:
                break
            if r2(p1,p2):
                p2=p2+1
            Rv=Rv+dRv(a,b,p1+p2)
            tiempo=tiempo+1
    else:
        for l in range(p2):
            if Rv>2:
                break
            if r2(p1,p2):
                p2=p2+1
            Rv=Rv+dRv(a,b,p1+p2)
            tiempo=tiempo+1
        for j in range(p1):
            if Rv>2:
                break
            if r1(p1,p2):
                p1=p1+1
            Rv=Rv+dRv(a,b,p1+p2)
            tiempo=tiempo+1
    #print (l,j)
#print(Rv,dRv(a,b,p1+p2))

return p1,p2, tiempo

```

*#Metodo que toma como parametro los plasmidos iniciales de cada tipo,
 su parametros de reproduccion(k1 o k2, h,#rondas)
 #y devuelve el numero de plasmidos finales de cada tipo(2 tipos)*

```
In [7]: def division(n):  
        c=cumbin(n)  
        r=random.random()  
        minimo=1  
        imin=n  
        for i in range(n):  
            if fabs(c[i]-r)<minimo:  
                imin=i  
                minimo=fabs(c[i]-r)  
        return imin,n-imin  
#Metodo que divide en dos un numero siguiendo un distribucion binomia  
l cpn p=0.5 dado un numero "n"
```

```

In [8]: from numpy import array
def bacterias(p1,p2,h,k1,k2,a,b,generaciones):

    tiempo=0
    rondas=2**generaciones
    gvacias=np.zeros(rondas)
    gtp1=np.zeros(rondas)
    gtp2=np.zeros(rondas)
    gtotalp1=np.zeros(rondas)
    gtotalp2=np.zeros(rondas)
    tiempo=0
    bact0=np.array([[p1,p2,tiempo]])

    bactf=np.array([])

    for j in range(rondas):
        # La primera bacteria (tiempo menor) es la que se reproduce
        bact0[0][0],bact0[0][1],tiempo=plasmidos(bact0[0][0],bact0[0]
[1],h,k1,k2,a,b,tiempo)
        #print(bact0[0][0],bact0[0][1],tiempo)
        # Los valores de las hijas
        p11,p12=division(bact0[0][0])
        p21,p22=division(bact0[0][1])
        # Las hijas en bactf
        bactf = np.array([[p11,p21,tiempo]])
        bactf = np.append(bactf,np.array([[p12,p22,tiempo]]),axis=0)
        # El padre es reemplazado por sus hijas
        bact0=bact0[1:]
        #print("-----antes-----")
        #print(bact0)
        #print("-----")
        bact0=np.append(bact0,bactf,axis=0)
        # Se ordena la poblacion con respecto al tiempo
        bact0 = np.sort(bact0.view('i8,i8,i8'), order=['f2'],
axis=0).view(np.int)
        #print("-----final-----")
        #print (bact0)
        #print("-----")

```



```
bactvacias=0.0
bactp1=0.0
bactp2=0.0
totalp1=0.0
totalp2=0.0
for i in range(len(bact0)):
    if bact0[i][0]==0:
        if bact0[i][1]==0:
            bactvacias+=1
        if bact0[i][1]!=0:
            bactp2+=1
    if bact0[i][1]==0:
        if bact0[i][0]!=0:
            bactp1+=1
    totalp1+=bact0[i][0]
    totalp2+=bact0[i][1]

gvacias=bactvacias
gtp1=bactp1
gtp2=bactp2
gtotalp1=totalp1
gtotalp2=totalp2

return gtp1, gtp2, gvacias, gtotalp1, gtotalp2
#Simulacion de plasmidos en una bacteria que se reproduce empezando c
on una sola bacteria con p1,p2 plasmidos
```

```

In [9]: def gen(p1,p2,i,k1,k2,a,b,gen,veces):

    tctp1=0
    tctp2=0
    tgvacias=0
    tgtotalp1=0
    tgtotalp2=0
    for r in range(veces):

        gtp1, gtp2, gvacias, gtotalp1,
gtotalp2=bacterias(p1,p2,i,k1,k2,a,b,gen)
        tctp1+=gtp1
        tctp2+=gtp2
        tgvacias+=gvacias
        tgtotalp1+= gtotalp1
        tgtotalp2+=gtotalp2

    #i=range(generaciones)

    logtotalgtp1=log10(tctp1/veces+2.0)
    logtotalgtp2=log10(tctp2/veces+2.0)
    logtotalvacias=log10(tgvacias/veces+2.0)

    razonfinal=logtotalgtp2/logtotalgtp1
    return razonfinal
#Metodo que repite la misma simulacion un "veces" numero de veces

    #plt.plot(i,logtotalgtp1,label="Dominante p1")
    #plt.legend()

    #plt.plot(i,logtotalgtp2,label="Dominantes p2")
    #plt.legend()

    #plt.plot(i,logtotalvacias,label="Vacias")
    #plt.legend()
    #plt.xlabel("Generaciones",size=20)
    #plt.ylabel("Celulas donde dominan los plasmidos",size=20)
    #plt.show()
    #plt.plot(i,logtotalgtp2/logtotalgtp1,label="Dom p2/Dom p1")
    #plt.plot(i,logtotalvacias/logtotalgtp1,label="Vacias/Dom p1 ")
    #plt.legend()
    #plt.xlabel("Generaciones",size=20)
    #plt.ylabel("Razon dominante 2/1",size=20)
    #plt.show()
    #plt.plot(logtotalgtp1,logtotalgtp2,label="Dom 2 vs dom 1")
    #plt.plot(logtotalgtp1,logtotalvacias,label="Dom vacias vs dom
1")
    #plt.legend()
    #plt.plot( xdata, yfit, '--k')
    #plt.xlabel("Dom 1",size=20)
    #plt.ylabel("Dom",size=20)
    #plt.show()

```

```

In [12]: def barrido(p1,p2,i,k2,a,b,generaciones,veces,n):
    start_time = time.time()
    k1=1.5*k2
    razon=np.zeros(n)
    num_cores=mp.cpu_count()
    hs=np.linspace(i,10*i,n)
    #for j in range(n):

        #razon[j]=gen(p1, p2, i, k1, k2, a,bs[j], generaciones, veces)
    razon=Parallel(n_jobs=num_cores)(delayed(gen)(p1, p2, h, k1, k2,
a,b, generaciones, veces) for h in hs)
    m,b=np.polyfit(hs,razon,1)
    y=np.zeros(n)
    for i in range(n):
        y[i]=b+m*hs[i]
    fig = plt.figure()
    plt.plot(hs/i,razon,label="Dom 2 vs dom 1")
    plt.plot(hs/i,y,label="reg")

    plt.title("p1="+ str(p1)+ "p2=" +str (p2)
+"i="+str(i)+"k2="+str(k2)+"generaciones=" + str (generaciones)+ "vec
es="+ str (veces)+"numcore="+ str(num_cores)+
"tiempo="+str(time.time() - start_time))
    plt.xlabel("h/i",size=20)
    plt.ylabel("Razon p2/p1",size=20)
    plt.show()
    fig.savefig("p1="+ str(p1)+ "p2=" +str (p2) +"i="+str(i)+"k2="+st
r(k2)+"generaciones=" + str (generaciones)+ "veces="+ str (veces)+"nu
mcore="+ str(num_cores)+
"tiempo="+str(time.time() - start_time)+".png")
    print (str(time.time() - start_time)+" segundos")
#Metodo que corre la simulacion en un rango de valores para k1 cierto
numero de veces

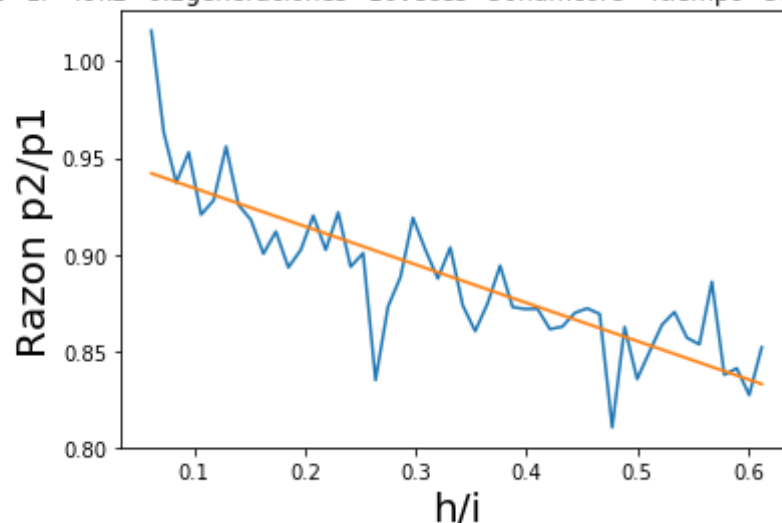
```

```

In [13]: barrido(10,1,3,0.2,1,100,10,50,50)

```

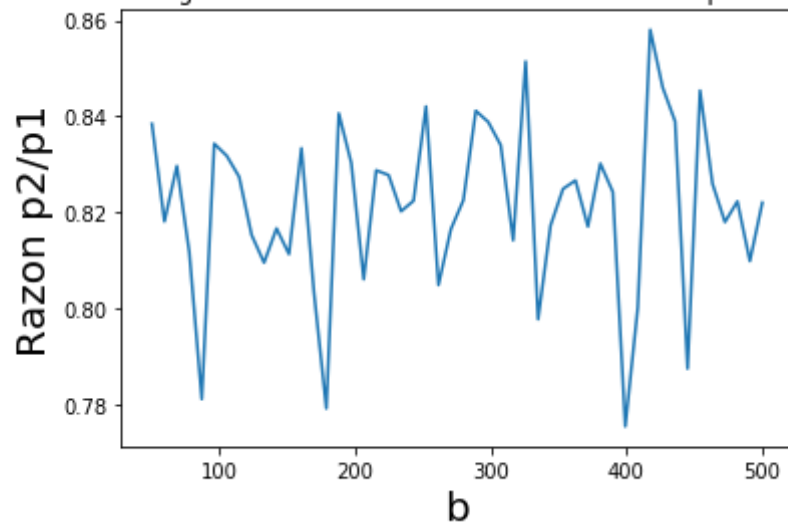
p1=10p2=1i=49k2=0.2generaciones=10veces=50numcore=4tiempo=5124.003643751144



5124.5252385139465 segundos

In [30]: `barrido(10,1,5,0.2,1,100,10,50,50)`

`p1=10p2=1i=5k2=0.2generaciones=10veces=50numcore=4tiempo=5694.624307632446`



5694.908969640732 segundos

```
In [31]: p1=1
p2=3
tiempo=0
bact0=np.array([[p1,p2,tiempo]])
print (bact0[0][0],bact0[0][1],bact0[0][2])
```

1 3 0

In []:

In []:

In []:

Bibliografía

- [1] Martin A. Nowak. Five rules of the evolution of cooperation. *National Institutes of health*, 2006.
- [2] Johan Paulsson y Mans Ehrenberg. Noise in a minimal regulatory network: plasmid copy number control. *Biophysics review*, 2001.
- [3] Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.
- [4] J. Walker R. Isaac and W. Williams. Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups. *Journal of Public Economics*, 1994.
- [5] Giraldo R. Ruiz M. Espinosa M. Solar, G. and Diaz Ramon. Replication and control of circular bacterial plasmids. *Microbiology and molecular biology reviews*, 1998.
- [6] Johan Paulsson. Multileveled selection on plasmid. *Genetic Society of America*, 2002.
- [7] Martin A. Nowak. Evolution of cooperation by multilevel selection. *PNAS*, 2006.
- [8] Bernhard Voelkl. Simulation of evolutionary dynamics in finite populations. *The Mathematica Journal*, 2011.
- [9] B. O. Bengtsson and H. Andersson. The population genetics of plant mitochondrial plasmids. *Theoretical Biology*, 1997.
- [10] M. Lipsitch Bergstrom, C. T. and B. R. Levin. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics journal*, 2000.
- [11] S. Bonhoeffer and M. A. Nowak. Intra-host versus inter-host selection—viral strategies of immune function impairment. 1994.
- [12] H. Bremer and S. Lin-Chao. Analysis of the physiological control of replication of *colE1*-type plasmid. 1986.

- [13] M. Brenner and J. Tomizawa. Quantitation of cole1- encoded replication elements. 1991.
- [14] L Chao. Levels of selection, evolution of sex in rna viruses, and the origin of life. 1991.
- [15] L Chao. The population genetics of rna viruses. 1994.
- [16] T. F. Cooper and J. A. Heinemann. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. 2000.
- [17] Ernst Fehr and Andreas Leibbrandt. A field study on cooperativeness and impatience in the tragedy of the commons. 2011.