



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di
Scienze Matematiche,
Fisiche e Naturali

Corso di Laurea Magistrale in
Biotecnologie Molecolari

Modellazione matematica di una comunità microbica stabile per lo studio della resistenza antimicrobica mediata da plasmide

Candidata
Elisa Sestini

Relatore
Prof. Marco Fondi

Correlatore
Prof. Alessio Mengoni

Anno Accademico 2025/2026

Indice

1	Introduzione	3
1.1	Comunità microbiche	3
1.1.1	Tipi di relazioni	3
1.1.2	Comunità microbiche sintetiche	4
1.2	Antibiotico-resistenza	5
1.3	Modelli matematici	7
1.4	Tipi di modelli	7
1.4.1	Modelli ecologici	9
1.4.2	Modello Lotka-Volterra	9
1.4.3	Modelli deterministici e stocastici	12
2	Scopo della ricerca	13
2.1	Approccio	13
3	Materiali e Metodi	14
3.1	R	14
3.2	MATLAB	14
3.2.1	Fitting sui dati sperimentali	15
3.2.2	Intervalli di confidenza	18
3.2.3	Regressione	18
3.3	Dati sperimentali	20
3.3.1	Chi.Bio	20
3.3.2	Plate counting	21
3.3.3	Citometria a flusso	21
3.4	Comunità microbica sintetica	21
3.4.1	Antibiotici utilizzati	22
3.4.2	Tipi di comunità per lo studio di antibiotico-resistenza con ampicillina e colistina	23
4	Risultati e discussione	25
4.1	Descrizione dei quattro esperimenti	25
4.2	Analisi dati gli esperimenti	26
4.3	Analisi della crescita dei ceppi sui terreni	33
4.4	Analisi dati sul secondo esperimento	34
4.5	Modello matematico delle comunità in assenza di antibiotico	38
4.5.1	Fitting del modello utilizzando le equazioni differenziali	39
4.5.2	Misura dell'intervallo di confidenza al 95%	42
4.5.3	Simulazione del modello nel lungo periodo e confronto con i dati sperimentali	43
4.6	Modello matematico delle comunità in presenza di antibiotico	45
4.7	Studio della degradazione di ampicillina e colistina	47
4.8	Studio del trasferimento del plasmide di antibiotico-resistenza	49

4.9	Costruzione del modello completo	52
4.10	Studio del cambiamento dei parametri ε e η tramite grid map e heatmap .	53
4.11	Difficoltà riscontrate durante la costruzione del modello	56
5	Conclusioni	58
5.1	Prospettive future e sviluppo del progetto	58
6	Materiale supplementare	60
6.1	Metodi alternativi di fitting	60
6.2	Modelli alternativi per la comunità senza antibiotico	60
6.3	Modelli per prevedere i dati	61
6.4	Modifiche della simulazione del modello	61
6.5	Simulazione degli aminoacidi	62
6.6	Script per calcolare gli intervalli di confidenza al 95%	62
6.7	Script del modello completo	64
	Bibliografia	76

1 Introduzione

1.1 Comunità microbiche

Una comunità microbica è un insieme di microrganismi che vivono nello stesso ambiente e interagiscono tra loro. L'unione dei microrganismi all'interno della comunità offre vantaggi notevoli per la sopravvivenza delle singole specie, che ottimizzano il consumo delle risorse, si proteggono dalle minacce esterne e comunicano tra loro.

La comunità ha delle proprietà macroscopiche che emergono dalle interazioni microscopiche delle specie che la formano. Queste proprietà emergenti sono i pattern o le funzioni che non possono essere dedotti dalla semplice somma delle proprietà dei singoli costituenti. Esempi di tali caratteristiche emergenti sono la resilienza alle perturbazioni biotiche e abiotiche, la coesistenza stabile e le capacità metaboliche o biochimiche collettive. Le proprietà emergono quando sono raggiunte specifiche soglie di grandezza e connettività della comunità, anche se tali soglie sono difficili da identificare a causa della natura non lineare di questi tratti collettivi. Ogni soglia è specifica per ciascuna proprietà e può dipendere da fattori ambientali [1].

Oltre al ruolo che hanno nella generazione delle proprietà emergenti, le interazioni tra le specie sono uno dei fattori primari coinvolti nella modifica dell'ambiente circostante e nella determinazione di struttura, funzione e dinamica delle comunità microbiche. Le relazioni instaurate possono essere per contatto diretto con i microrganismi vicini oppure attraverso metaboliti e molecole di segnalazione chimica diffusibili a distanza. Tali relazioni, inoltre, possono essere vantaggiose oppure deleterie, obbligate oppure facoltative [2]. Le principali relazioni sono il mutualismo, la sintrofia, la competizione, la predazione, il parassitismo e il commensalismo.

1.1.1 Tipi di relazioni

Una delle interazioni positive (+/+) più comuni è quella in cui una o più specie si nutrono dei prodotti di altri membri della comunità e i microrganismi che le instaurano hanno vantaggi reciproci per la sopravvivenza e per la crescita. Nelle interazioni mutualistiche obbligate, come le sintrofie, le singole specie non possono sopravvivere in assenza dei loro partner. Per esempio, in un'interazione di cross-feeding, tutte le specie rilasciano nutrienti essenziali per la crescita dei loro partner e in assenza di una delle specie, la sopravvivenza di tutta la comunità risulta compromessa. In alternativa, ci sono casi di interazioni benefiche in cui una specie utilizza i prodotti di scarto di un'altra e in cambio mantiene le condizioni termodinamiche favorevoli all'altra, come nel caso dei batteri metanogeni e dei solfato-riduttori. I metanogeni utilizzano i sottoprodotti della fermentazione prodotti dai solfato-riduttori e in cambio rendono le condizioni termodinamiche favorevoli ai solfato-riduttori eliminando dall'ambiente di reazione l'idrogeno. Le interazioni benefiche obbligatorie possono essere un motivo per cui la maggior parte dei microrganismi non è coltivabile in laboratorio come monocultura, infatti potrebbero mancare le specie partner da cui dipendono per l'ottenimento di nutrienti o fattori di crescita essenziali [2].

Le relazioni antagoniste sono la competizione, il parassitismo e la predazione.

La competizione (-/-) avviene quando due o più specie batteriche si contendono le stesse risorse limitate, come i nutrienti. E' una delle interazioni più comuni e può portare all'esclusione di una specie dall'ambiente considerato. Rientrano in questa categoria anche i meccanismi di interferenza, tra cui l'inibizione della crescita di una specie tramite molecole tossiche prodotte da un'altra, come le batteriocine, peptidi antimicrobici attivi contro i batteri competitori, o l'uccisione diretta di una specie attraverso antibiotici prodotti da un'altra [2].

Il parassitismo (+/-) avviene per esempio tra batteri e batteriofagi ed è una relazione in cui il parassita trae un beneficio, mentre l'ospite viene danneggiato [2]. La predazione (+/-), invece, è una relazione in cui una specie si nutre direttamente di un'altra.

Infine, il commensalismo (+/0) è una relazione neutra, in cui una specie trae beneficio dall'interazione con un'altra, che non ne risente né positivamente né negativamente, come nel caso in cui un batterio utilizzi sottoprodotti di un altro senza apportarle né beneficio né danno.

Le relazioni microbiche positive o negative sono correlate alla stabilità della comunità. Una comunità è considerata stabile se le sue specie ritornano ad uno stato di equilibrio dopo una piccola perturbazione e la stabilità viene misurata attraverso la probabilità del ritorno allo stato iniziale, la dinamica di popolazione durante il ritorno e il tempo impiegato per il ritorno [3].

La stabilità della comunità è promossa dalla limitazione dei feedback positivi, dalla bassa diversità di specie nella comunità e dall'indebolimento delle interazioni ecologiche. Le relazioni cooperative tendono ad aumentare l'instabilità della comunità poiché causano feedback positivi e perché se l'abbondanza di una specie diminuisce, anche le abbondanze delle altre tendono a diminuire a causa della dipendenza dalla prima, allontanando il sistema dalla stabilità. Per questo motivo, le interazioni cooperative promuovono l'efficienza metabolica, ma diminuiscono la stabilità ecologica, che dipende da due meccanismi distinti. Il primo riguarda i feedback negativi introdotti dalle interazioni competitive, che smorzano quelli positivi, e il secondo è l'indebolimento delle interazioni ecologiche attraverso la ridondanza, ossia la sostituzione di interazioni cooperative forti con numerose interazioni più deboli [3].

1.1.2 Comunità microbiche sintetiche

Le comunità microbiche sintetiche sono modelli semplificati delle comunità microbiche naturali, che normalmente sono formate da decine, centinaia o migliaia di specie diverse. Possono essere ottenute coltivando insieme specie microbiche wild-type in un terreno di crescita simile all'habitat naturale, specie mutate geneticamente in modo targeted, come knockout genici, oppure specie non interagenti in condizioni ambientali particolari, che favoriscono la nascita di nuove interazioni [2]. Alcuni esempi di comunità microbiche sintetiche di *Escherichia coli* sono presenti nell'articolo di Mee et al. [4], in cui sono stati prodotti 14 mutanti per auxotrofia di *E. coli*, ognuno privo di un gene per la produzione di un aminoacido essenziale. I mutanti sono stati coltivati in coculture per studiare la crescita sintrofica degli auxotrofi e il cross-feeding metabolico, ossia lo scambio di metaboliti essenziali che governa la crescita e la composizione di comunità microbiche. Partendo da un modello di cross-feeding semplice, ossia quello di cocoltura a coppie, sono stati successivamente realizzati modelli di cocoltura a gruppi di tre e con tutti i mutanti insieme per ricostruire la complessità e la dinamica delle comunità naturali.

Dagli studi sulle comunità sintetiche, Mee et al. [4] hanno notato che gli aminoacidi biosinteticamente costosi, come metionina, lisina, isoleucina, arginina e gli aromatici, promuovono interazioni cooperative di cross-feeding più forti rispetto a quelli più economici. La biosintesi degli aminoacidi è ottimizzata per massimizzare l'efficienza energetica delle specie in due modi contemporaneamente: gli aminoacidi meno costosi sono più utilizzati rispetto agli altri; gli aminoacidi più costosi sono prodotti soltanto da una piccola frazione di membri in una comunità e sono resi disponibili agli altri tramite cross-feeding, favorendo la crescita sinergica. La generazione e l'evoluzione di relazioni cooperative per la produzione di aminoacidi costosi o per altre funzioni energeticamente sfavorevoli è sostenuta anche dalla *Black Queen Hypothesis* (BQH), che assume che la cooperazione tra le specie emerga per tratti puramente egoistici e che i membri della comunità che svolgono funzioni costose e indesiderate rilascino parte dei loro prodotti per renderli utilizzabili ad altre specie. I membri che non svolgono tali funzioni vanno incontro a una perdita adattativa dei geni corrispondenti e questo genera la formazione di "helpers" e "beneficiaries" nella comunità, che hanno associazioni obbligatorie. Il termine *black queen* si riferisce a queste funzioni, evitate dalla maggior parte delle specie, che portano alla generazione di cross-feeding massimizzando la produzione dei prodotti vitali per tutta la comunità a partire da membri helper. E' stato dimostrato che l'evoluzione di scambi cooperativi riduce la produttività della comunità e questo tipo di comportamento cooperativo evolve solo sotto specifici regimi demografici caratterizzati da mixing genetico intermedio [2].

Le coppie di mutanti che condividono intermedi comuni dei pathway di reazione hanno una crescita sinergica, **tuttavia quando si scala a modelli di cocoltura a tre e a quattordici membri**, la sinergia non è prevedibile, dal momento che possono essere presenti epistasi positive e negative. Nelle comunità sintetiche di tre membri, in cui ognuno è auxotrofo per due aminoacidi, la crescita di ogni membro avviene soltanto se entrambi gli aminoacidi sono forniti dai partner, ossia se sono presenti contemporaneamente tutti i membri, e le interazioni cooperative maggiori sono presenti tra ceppi con auxotrofie per metionina, fenilalanina e lisina. Questo studio ha messo in luce che gli approcci modellistici sono fondamentali per descrivere la complessità della dinamica delle comunità microbiche [4].

1.2 Antibiotico-resistenza

Le comunità microbiche possono essere disturbate da diversi fattori esterni, tra cui i trattamenti antibiotici. Questi farmaci disturbano l'intera comunità, cambiandone la composizione e le funzioni metaboliche e riducendo l'abbondanza e la diversità microbiche. Il loro effetto, inoltre, indebolisce la resistenza della comunità, ossia la capacità delle colonie residenti di escludere ceppi invasivi attraverso la competizione dei nutrienti, la produzione di batteriocine e l'immunità mediata dall'ospite. Una volta indebolita la resistenza della comunità, si aprono nicchie ecologiche che favoriscono l'invasione di ceppi opportunistici o resistenti [5].

Negli studi sull'antibiotico-resistenza, trattare i patogeni come isolati limita la predizione dei risultati perché non dà informazioni su come le comunità rispondono alle perturbazioni, e colpire un patogeno specifico può favorire i competitori resistenti a causa dell'alterazione della struttura della comunità. Inoltre, le interazioni cooperative tra ceppi resistenti e sensibili possono condurre a una resistenza a livello della comunità, come nel caso della degradazione degli antibiotici da parte delle β -lattamasi prodotte dai batteri resistenti, che crea una nicchia extracellulare protettiva per i ceppi sensibili. Per questo motivo, è importante studiare l'antibiotico-resistenza a livello della comunità [5].

Il microbioma, ossia la comunità di microrganismi e il loro materiale genetico complessivo, è un sistema ecologico che influenza attivamente come l'antibiotico-resistenza emerge, persiste e si diffonde tra ospiti e ambienti. La perdita dei ceppi suscettibili dovuta alla presenza dell'antibiotico aumenta la disponibilità delle risorse e la frequenza delle interazioni dirette tra le cellule batteriche, accelerando il trasferimento dei geni di resistenza agli antibiotici. I microbiomi con maggiore diversità recuperano più velocemente perché la diversità promuove la costruzione di nicchie e la ridondanza funzionale. Per questo motivo, la selezione batterica è dovuta alla struttura della comunità e non solo all'intensità e alla durata dell'esposizione all'antibiotico [5].

in italico

L'antibiotico-resistenza si può dividere in tre processi dinamici: emergenza, diffusione e persistenza. L'emergenza avviene per mutazione **de novo**, trasferimento genico orizzontale (HGT, horizontal gene transfer) e arrivo di individui resistenti provenienti da un altro ambiente. Questi eventi stocastici dipendono dalla dimensione della popolazione, dal tasso spontaneo con cui emergono nuove varianti (per mutazione, **trasferimento o immigrazione**) e dalla frazione di tali eventi che conferisce resistenza; questi fattori possono essere alterati dalle interazioni con il microbioma circostante. La diffusione dipende dalla deriva genetica e successivamente dalla selezione, che dipende a sua volta dall'ambiente e dagli effetti dati dalla mutazione. Infine, la **persistenza** avviene anche in mancanza della pressione selettiva degli antibiotici per tre motivi. Il primo è che alcune mutazioni conferiscono un vantaggio anche in presenza di altri fattori di stress, il secondo è che l'arrivo continuo di ceppi resistenti tramite mutazione o immigrazione mantiene le mutazioni costose a frequenze intermedie e il terzo è che le mutazioni compensatorie permettono di superare il costo di fitness associato alla resistenza [5].

con persistenza si intende solitamente un altro processo (non genetico) che permette di resistere all'azione di un antibiotico. Cambierei quindi con mantenimento anziché persistenza

La resistenza a un antibiotico spesso persiste a lungo anche quando non è più presente l'antibiotico, rimanendo per mesi o anni, sostenuta da meccanismi come l'occupazione di nicchie favorevoli, le mutazioni compensatorie, che riducono il costo di fitness della resistenza, e il trasferimento genico orizzontale nel microbioma.

Il trasferimento genico orizzontale diffonde la resistenza tramite gli MGE (mobile genetic elements), come plasmidi, ossia molecole di DNA autoreplicanti, trasposoni, cioè frammenti di DNA che si muovono all'interno e tra i genomi, e integroni, ossia piattaforme genetiche che catturano ed esprimono cassette geniche. L'efficienza del trasferimento dipende dalla densità della popolazione, dalle reti di contatto microbico e dalla compatibilità tra i donatori e i riceventi. Brevi esposizioni agli antibiotici possono aumentare l'abbondanza degli MGE, permettendo a questi elementi di trasferirsi a uno spettro più ampio di ospiti microbici. Poiché questi elementi mobili possono persistere e diffondersi anche dopo la rimozione della pressione antibiotica, l'insieme dei geni di resistenza presenti nella comunità spesso non torna al proprio stato precedente. Le mutazioni che migliorano il controllo sulla replicazione dei plasmidi, l'efficienza di coniugazione e l'integrazione nei cromosomi contribuiscono alla stabilizzazione delle associazioni tra plasmide e ospite [5]. Anche l'organizzazione spaziale delle comunità microbiche modula il flusso genico. Negli ambienti in cui le popolazioni microbiche sono più dense si forma una struttura a livello microscopico che guida i contatti batterici e la diffusione delle molecole. Questa aggregazione può favorire la coniugazione tra donatori e riceventi, mentre la compartimentazione la limita. Di conseguenza, l'eterogeneità spaziale può creare differenze nella diffusione della resistenza a seconda della zona [5].

Non è ancora chiaro se la strategia migliore per contrastare le infezioni batteriche sia l'utilizzo di dosi elevate di antibiotici, o se queste possano favorire lo sviluppo di resistenza.

Il modo in cui vengono somministrati gli antibiotici incide infatti sia sull'antibiotico-resistenza sia sulla tolleranza che si sviluppano nella comunità microbica. Il livello e la durata della perturbazione indotta dipendono dallo spettro dell'antibiotico, dal dosaggio, dalla lunghezza del trattamento e dalla struttura del microbiota residente; i trattamenti di breve durata sono spesso più efficaci nel ridurre la resistenza senza compromettere il recupero dei pazienti [5]. Se invece il trattamento non ha la durata o il dosaggio adeguati, si sviluppa la tolleranza agli antibiotici, ossia la capacità di una popolazione sensibile di sopravvivere alla presenza dell'antibiotico riducendo il suo tasso di crescita e di morte senza modifiche genetiche. L'emergenza e il mantenimento della tolleranza sono determinati dal tipo e dalla forza delle interazioni microbiche nella comunità in presenza dell'antibiotico: interazioni ^{in italico} più forti tendono a eliminare dalla comunità la specie patogena che genera cellule **persistenter**, impedendole di resuscitare, mentre interazioni più deboli le permettono di persistere e ripopolare la comunità quando la pressione antibiotica cessa [6]. Per questo motivo, la competizione naturale tra batteri resistenti e sensibili influenza la strategia di trattamento da utilizzare e comunità microbiche diverse possono rispondere agli antibiotici in modo diverso a parità di dose.

1.3 Modelli matematici

I modelli matematici permettono di esplorare come interagiscono i fattori biologici, interpretare i processi biologici multiscala e predire le loro conseguenze sulla dinamica di antibiotico-resistenza. La complessità degli ecosistemi microbici rende i modelli matematici indispensabili, poiché offrono un modo per semplificare, simulare e predire la dinamica, che è difficile da studiare soltanto sperimentalmente. I modelli matematici contribuiscono allo studio dell'antibiotico-resistenza in due modi. Nel primo caso, permettono di trarre inferenze basate sul modello per identificare i processi biologici che danno vita a pattern di resistenza osservati e testare se il modello cattura i processi biologici reali, così da dimostrare se è realistico e interpretabile. Nel secondo caso, i modelli permettono di prevedere la resistenza e dare priorità a un trattamento strategico sia a livello di individuo che di popolazione. Dal punto di vista individuale, possono guidare la scelta personalizzata di un antibiotico per un paziente specifico, mentre a livello di popolazione il modello può predire in che misura cambiano l'incidenza e la prevalenza dell'antibiotico-resistenza adottando specifiche politiche e linee guida [5].

I modelli matematici predittivi aiutano a decifrare come le interazioni tra le specie microbiche governano la dinamica e l'evoluzione delle comunità [1] e possono avere un grande valore terapeutico poiché guidano la selezione degli antibiotici per ridurre il rischio di malattie indotte dall'antibiotico [7]. Tali modelli possono essere sviluppati a partire da una conoscenza quantitativa delle interazioni microbiche basata sui dati ecologici di serie temporali, che mostrano la composizione di una comunità in un punto temporale discreto, e hanno potere predittivo e analitico poiché offrono predizioni sulla struttura della comunità e sulla risposta agli stimoli esterni [7]. Lavorare con le serie temporali permette di capire non solo lo stato attuale della comunità microbica, ma anche la sua dinamica nel tempo [8].

1.4 Tipi di modelli

Esistono più tipi possibili di modelli matematici e la scelta del modello dipende dal tipo di ricerca richiesto e dai dati disponibili per stimare i parametri da inserire all'interno

del modello. Il tipo di ricerca riguarda quale tipo di conoscenza si vuole ricavare dal modello; questa può essere una comprensione generale del fenomeno di interesse oppure una misura più specifica per generare predizioni a partire dal fitting sulle osservazioni sperimentali. Riguardo invece ai dati disponibili per la stima dei parametri del modello, quando sono presenti le misurazioni sperimentali, viene utilizzato un approccio guidato dai dati, in cui i parametri vengono stimati tramite fitting su di esse. Quando invece i dati sperimentali sono limitati, il modello può essere costruito utilizzando l'approccio dei principi primi (*first principles*), ossia partendo dalle conoscenze generali del sistema, come le reazioni metaboliche codificate da un genoma, senza la necessità di calibrazione sui dati sperimentali. Questo tipo di modelli permette di testare la risposta del sistema alle perturbazioni inaccessibili agli esperimenti e alle osservazioni. In generale, rispetto alle analisi statistiche, i modelli matematici aiutano a generare ipotesi meccanicistiche e a stabilire relazioni di causalità [1].

La qualità dei modelli dipende dall'accuratezza e dall'incertezza nella stima dei parametri, per cui la scelta del tipo di modello e della sua complessità è guidata dalla disponibilità dei dati. I parametri richiesti possono essere categorizzati in parametri a livello dei membri, come la *carrying capacity*, e in parametri di interazione, come i tassi di *cross-feeding*. I parametri dovrebbero essere ottenuti nel contesto delle comunità, ma è difficile a causa delle difficoltà sperimentali, per cui molti modelli ricorrono ai parametri ottenuti da monoculture o studi di comunità in piccola scala. Lo sviluppo dei metodi omici e delle colture *high-throughput* può fornire informazioni per superare questa difficoltà, infatti per esempio le analisi di metagenomica possono suggerire quali membri della comunità interagiscono tramite covariazione statistica e aiutare a stimare i tassi di crescita in situ usando la *read coverage* sull'origine di replicazione. I dati di metatrascrittomica e metabolomica invece possono essere usati per vincolare la struttura del modello selezionando le interazioni consistenti con i dati [1].

I modelli matematici, inoltre, sono indispensabili per stabilire le connessioni quantitative tra la struttura della comunità, intesa come composizione e connettività dei suoi componenti, e le funzioni emergenti, data la loro natura non lineare. Poiché le dinamiche e le proprietà emergenti della comunità microbica possono essere attribuite all'interazione di vari network complessi, per scegliere l'approccio di modellazione migliore, che le predica in modo accurato, è necessario identificare quali network debbano essere inclusi. Ci sono due approcci per modellare la dinamica di comunità e le proprietà emergenti. Il primo sono i modelli ecologici, in cui le cellule o le specie sono le unità di base e le interazioni tra loro sono i punti centrali; il secondo sono i modelli metabolici a scala genomica (*genome-scale metabolic models*), in cui le reazioni intracellulari sono le unità di base e la generazione o il consumo di nutrienti sono i punti centrali. Comparare i modelli sulla base delle predizioni di proprietà emergenti è problematico a causa della variazione nei dettagli biochimici e nel numero di dimensioni modellate [1].

Gli approcci ecologici sono quattro: Lotka-Volterra, MacArthur consumer-resource, trait-based e individual-based. Insieme ai modelli metabolici *genome-scale* possono essere suddivisi in approcci top-down e bottom-up. Gli approcci top-down sono utilizzati nei modelli Lotka-Volterra, MacArthur consumer-resource e trait-based e sono più adatti a modellare resilienza della comunità e coesistenza, simulando l'adattamento a livello di popolazione e non catturando le dinamiche micro-evoluzionistiche a partire dalle mutazioni nelle singole cellule. Gli approcci bottom-up, invece, sono utilizzati nei modelli individual-based e metabolici *genome-scale* e sono adatti a modellare i range di tratti fenotipici e i pattern

spaziali, in cui la dinamica e la funzione del sistema intero sono governate da quella delle cellule individuali [1].

In base alle caratteristiche temporali e spaziali della proprietà di interesse, è possibile combinare gli approcci, poiché ognuno ha vantaggi e limitazioni e la distinzione tra i modelli non è categorica, quindi l'adattamento di un certo modello può fargli cambiare categoria. Una sinergia di approcci top-down e bottom-up può essere utile a simulare e predire al meglio una certa proprietà [1] e analogamente può favorire la progettazione di consorzi sintetici [2]. I sistemi dinamici, come le comunità microbiche, sono processi tempo-dipendenti descritti da ODE (equazioni differenziali ordinarie), per cui più tipi di modelli possono essere utilizzati in base ai dati disponibili per descrivere le interazioni ecologiche tra i membri della comunità, che possono essere dirette, se un organismo ha effetto diretto sulla fitness dell'altro, come nel modello Lotka-Volterra generalizzato, o indirette, se un organismo utilizza i composti metabolici prodotti da un altro, descritte dai modelli consumer-resource.

1.4.1 Modelli ecologici

I modelli ecologici per le comunità microbiche sono stati adattati da quelli classici utilizzati negli ecosistemi di macrofauna, considerando tre differenze principali. La prima riguarda il grande numero di individui delle comunità microbiche; la seconda riguarda i tempi di generazione brevi, che insieme al trasferimento genico orizzontale portano dinamiche ed evoluzione rapide delle interazioni, alterando processi come la migrazione e l'esclusione dalla competizione; la terza riguarda la separazione del layer trofico, poiché le comunità microbiche trasferiscono la biomassa sia attraverso interazioni trofiche, come quelle preda-predatore, che soprattutto tramite molecole secrete durante il cross-feeding. I vari modelli matematici sono indirizzati a diverse proprietà ecologiche: stabilità della composizione, risposta alle perturbazioni abiotiche o alle invasioni, dipendenza dall'ambiente delle interazioni tra le specie, pattern spazio-temporali di colonizzazione e successione e conflitto tra interessi individuali e collettivi, richiamato ad esempio dalla *Black Queen Hypothesis* [1].

1.4.2 Modello Lotka-Volterra comune

Lotka-Volterra è il modello ecologico più **basico** [1], utilizzato in origine per descrivere le dinamiche di una popolazione interagente in un ecosistema, come nel caso delle interazioni preda-predatore, e successivamente per connettere le interazioni tra le specie alla dinamica di comunità complesse. E' formato da un set di equazioni differenziali di primo grado, accoppiate e non lineari, che permettono di modellare la dinamica di comunità e di descrivere l'evoluzione temporale dell'abbondanza delle specie microbiche [2].

Il modello Lotka-Volterra classico è utilizzato per studiare le interazioni tra due specie e può essere risolto numericamente o analiticamente. I suoi limiti principali, tuttavia, sono due. Il primo consiste nella generazione di oscillazioni neutralmente stabili, ossia che non hanno un periodo o un'ampiezza caratteristica, al posto di oscillazioni cicliche; questo causa risultati diversi qualitativamente a partire da piccole modifiche nel modello. Il secondo limite consiste nel considerare interazioni soltanto tra due specie [1]. Questo è superato nel modello Lotka-Volterra generalizzato (Figura 1.1), che è un sistema di equazioni differenziali ordinarie che descrive le dinamiche del microbioma aggiungendo un numero non limitato di specie interagenti, i loro tassi di crescita e le loro interazioni a coppie [5].

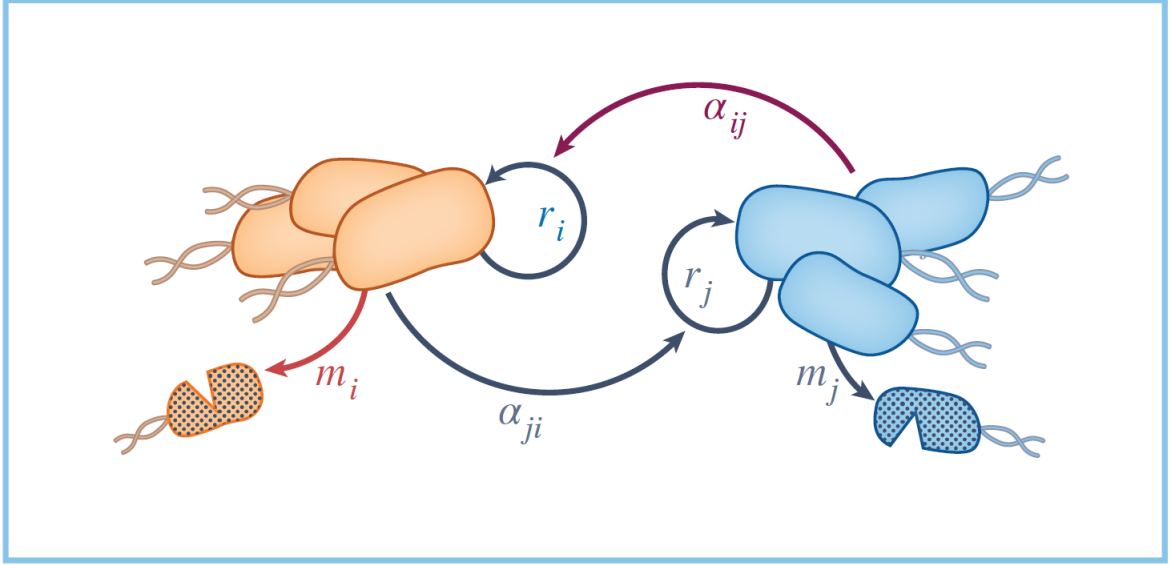


Figura 1.1: Schema del modello Lotka-Volterra generalizzato, con più specie batteriche che interagiscono tra loro. Immagine presa dall'articolo [1].

Il modello Lotka-Volterra generalizzato utilizza pochi parametri, che possono essere dedotti dai dati di co-coltura, dai dati di sopravvivenza allo stato stazionario o dalle correlazioni osservate nei dataset metagenomici, e può essere descritto da equazioni differenziali 1.1 che descrivono la crescita della popolazione:

$$\frac{dB_i}{dt} = B_i * ((r_i + \sum_N^j \alpha_{ij} B_j) - m_i) \quad (1.1)$$

in cui B_i rappresenta la concentrazione o il numero di cellule della specie i , $\frac{dB_i}{dt}$ il tasso di variazione della specie i nel tempo, r_i il tasso di crescita intrinseco della specie i , α_{ij} il coefficiente di interazione tra la specie i e la specie j , B_j la concentrazione o il numero di cellule della specie j e m_i il tasso di mortalità della specie i . I coefficienti di interazione interspecie α possono essere modificati per mostrare diversi tipi di interazioni attraverso il cambiamento del segno, che indica effetti positivi o negativi di una specie sull'altra, come nel caso di cooperazione o competizione, e il loro ordine di grandezza indica la forza dell'interazione. Il tasso di crescita di una popolazione è visto come l'insieme del tasso di crescita intrinseco della specie e degli effetti lineari esercitati dalle altre popolazioni, mentre la crescita di ogni popolazione è modulata dagli effetti, positivi o negativi, esercitati dalle interazioni con le altre specie, ed è ridotta dal proprio tasso di mortalità [1].

Una delle applicazioni del modello Lotka-Volterra generalizzato consiste nel predire la dinamica temporale, la stabilità e la resilienza di comunità microbiche complesse in varie condizioni sperimentali, come nel caso in cui vengano aggiunte perturbazioni esterne dipendenti dal tempo. Stein et al. [7] hanno utilizzato un modello Lotka-Volterra generalizzato aggiungendo una perturbazione esterna tempo-dipendente per quantificare i tassi di crescita delle specie microbiche modellate, le interazioni tra le specie e la suscettibilità dei gruppi microbici alle perturbazioni esterne tempo-variabili, come l'esposizione agli antibiotici o i cambi nella disponibilità dei nutrienti. Hanno utilizzato i parametri ottenuti

da dati di serie temporali per predire numericamente la dinamica del microbiota e **pr** calcolarne gli steady state, ossia gli stati stazionari in cui il sistema non varia più nel tempo. Poiché le equazioni del modello non sono lineari, possono essere presenti più steady state; la stabilità di ciascuno di essi, che determina il comportamento a lungo termine delle traiettorie del sistema, è valutabile analiticamente.

I vantaggi del modello Lotka Volterra sono due **principali**. Il primo consiste nella semplicità del sistema: il basso numero di parametri necessari, che possono essere stimati a partire da dati di serie temporali di esperimenti in vivo e in vitro, rende il modello facile da generalizzare [1]. La sua flessibilità, inoltre, permette di combinare interazioni competitive e cooperative [2]. Il modello generalizzato, in particolare, può accomodare un numero arbitrario di specie [1], ed è stato applicato sia a grandi comunità multi-specie, come quelle intestinali o presenti nel formaggio, sia a sistemi più semplici a due popolazioni, come le interazioni batteri-batteriofagi [2]. Il secondo vantaggio consiste nella possibilità di identificare analiticamente le condizioni di equilibrio e la stabilità [1].

i seguenti

I limiti principali presenti nel modello, invece, sono **cinque**.

Il primo limite consiste nel fatto che le interazioni sono statiche nel tempo e nello spazio, mentre nei sistemi naturali i cambi nell'ambiente abiotico o nelle densità delle specie possono modificare le interazioni, per esempio passando dal mutualismo alla competizione. Questo può essere superato permettendo ai segni delle interazioni tra le specie di variare nel tempo attraverso una carrying capacity congiunta, che indica il numero massimo di individui che le risorse disponibili in un ambiente possono sostenere.

Il secondo limite riguarda i feedback positivi, che possono causare una crescita libera della popolazione, per cui il fitting sui dati di questo modello può sottostimare la prevalenza del mutualismo. Questo limite può essere superato aggiungendo più parametri, poiché la crescita libera può essere bloccata sostituendo le interazioni lineari tra le specie con funzioni saturanti non lineari.

La terza limitazione riguarda la descrizione della comunità soltanto attraverso interazioni a coppie, anche se è possibile aggiungere interazioni di ordine superiore. Ciascuna di queste aggiunte, tuttavia, comporta la difficoltà di un'interpretazione significativa dei parametri e il numero di coefficienti di interazione scala in modo quadratico rispetto ai membri della comunità. Per considerare le interazioni di ordine superiore senza aumentare troppo lo spazio dei parametri, i coefficienti di interazione a coppie possono essere ridefiniti in modo da incorporare anche l'effetto delle interazioni di ordine superiore.

Il quarto limite consiste nella mancata considerazione dei processi stocastici, come la deriva genetica o le fluttuazioni ambientali, ossia i cambi di pH, temperatura o livello dei nutrienti. Per superare tale limite, la stocasticità può essere incorporata nel modello aggiungendo termini di rumore alle equazioni e l'eterogeneità ambientale può essere modellata attraverso fluttuazioni nel tasso di morte e interazioni che variano nel tempo. Inoltre, gli intervalli di confidenza possono essere ottenuti tramite il fitting stocastico.

L'ultimo limite riguarda le interazioni mediate dalle piccole molecole, come lo scambio di metaboliti o il quorum sensing, che non sono modellate esplicitamente, limitando lo studio dell'emergenza di proprietà biochimiche a livello di comunità. Per superare questo limite, le interazioni indirette mediate dai metaboliti possono essere modellate esplicitamente con un modello di interazioni quadratiche specie-metaboliti, che permettono la coesistenza stabile e la diversità tramite il cross-feeding [1].

Per riassumere, il modello Lotka-Volterra generalizzato è un modello deterministico di dinamica ecologica tempo-dipendente [7] e viene utilizzato per studiare la resilienza e la

diversità della comunità, la coesistenza stabile delle specie [1] e l'impatto di perturbazioni esterne tempo-dipendenti [7]. Permette di identificare i membri di una comunità e il loro ruolo e si utilizza quando le misure dei tratti delle specie non sono disponibili [1].

1.4.3 Modelli deterministici e stocastici

I modelli matematici possono essere deterministici o stocastici. Un modello deterministico è utilizzato quando il sistema considerato è omogeneo e formato da un alto numero di componenti e produce sempre lo stesso risultato se i parametri iniziali sono gli stessi. Tra i modelli analizzati, Lotka-Volterra generalizzato e MacArthur consumer-resource sono deterministici [1], poiché la dinamica microbica considerata è predicibile e guidata da equazioni basate su parametri, come il tasso di crescita, i coefficienti di interazione e i tassi di utilizzo delle risorse, che sono tutti assunti come costanti o influenzati da parametri dell'ambiente esterno. Anche il modello metabolico genome-scale può essere considerato deterministico, poiché lo stato metabolico ottimizzato ottenuto attraverso FBA è lo stesso a partire dalle stesse condizioni.

Un modello stocastico, invece, incorpora la casualità e l'incertezza, riflettendo la variabilità delle comunità microbiche. Considera che le interazioni microbiche e i fattori ambientali non sono sempre predicibili e che gli eventi casuali possono influenzare la dinamica di comunità. Il modello individual-based è stocastico per natura, poiché tramite il suo approccio bottom-up modella ogni membro della comunità individualmente e cattura eterogeneità ambientale e interazioni locali. Le fluttuazioni stocastiche che questo modello può catturare sono la deriva genetica, le mutazioni, il trasferimento genico orizzontale e le interazioni cellula-cellula, che portano a proprietà emergenti come resilienza, coesistenza e complessità fenotipica [1].

E' possibile rendere i modelli deterministici stocastici. Il modello Lotka-Volterra generalizzato, per esempio, può essere reso stocastico in diversi modi [1] ed è possibile utilizzare sia l'approccio stocastico che quello deterministico per descrivere la dinamica di una comunità di tre batteri, utilizzata per studiare la resuscitazione dei persister [6]. In particolare, dal modello stocastico è stato notato che le specie patogene possono resuscitare dopo un trattamento con gli antibiotici e dominare la comunità, mentre con il modello deterministico non è mai stato ottenuto qualcosa di simile. Da questo si deduce che le fluttuazioni della popolazione possono avere un ruolo nella resilienza delle popolazioni dopo una perturbazione. Anche i modelli consumer-resource di MacArthur possono essere resi stocastici tramite l'incorporazione di rumore demografico [1].

Per riassumere, i modelli deterministici offrono semplicità computazionale e analitica, ma non riescono a rappresentare l'intera complessità degli ecosistemi microbici, poiché, per esempio nel caso dei modelli Lotka-Volterra generalizzati, non sono considerate la variabilità casuale e le interazioni di alto ordine oltre a quelle a coppie. I modelli stocastici, invece, catturano le proprietà emergenti, ma sono computazionalmente costosi e meno adatti a fare predizioni generalizzate, come nel caso dei modelli individual-based. La combinazione di entrambi gli approcci permette di descrivere comunità grandi, ma considerando la variabilità ed evitando costi computazionali eccessivi.

2 Scopo della ricerca

La ricerca si inserisce nell'ambito dello studio della resistenza agli antibiotici e il suo scopo è di realizzare un modello Lotka-Volterra generalizzato per descrivere la dinamica di una comunità batterica e il trasferimento genico orizzontale di plasmidi di antibiotico-resistenza. La comunità microbica stabile sotto studio è composta da tre ceppi auxotrofi di *Escherichia coli* che stabiliscono interazioni mutualistiche per garantire la sopravvivenza dell'intera comunità. In alcuni esperimenti, uno dei tre ceppi contiene un plasmide di antibiotico-resistenza e ha la capacità di trasferirlo alle specie sensibili tramite coniugazione, permettendo di studiare la propagazione della resistenza attraverso il trasferimento genico orizzontale. Questo set sperimentale permette di rispondere le seguenti domande:

- Qual è l'impatto dell'antibiotico in assenza di resistenza?
- Come cambia la dinamica in presenza del plasmide di antibiotico-resistenza e del suo trasferimento?
- Qual è il ruolo del tasso di coniugazione e della sensibilità all'antibiotico nella dinamica della comunità?

2.1 Approccio

Il modello viene realizzato seguendo un approccio data-driven, ossia i suoi parametri e i relativi intervalli di confidenza vengono ricavati a partire dal fitting sui dati sperimentali di serie temporali, che contano il numero di colonie dei singoli ceppi della comunità nel tempo. Lo scopo del modello consiste nel simulare la dinamica della comunità batterica in assenza e in presenza di antibiotico, e con o senza il trasferimento del plasmide da un ceppo agli altri. E' inoltre possibile simulare l'andamento della comunità a tempi successivi a quelli sperimentali, per comprendere il comportamento delle specie a lungo termine, e modificare i parametri riferiti all'effetto dell'antibiotico e della coniugazione del plasmide per osservare come cambia il comportamento della comunità e comprendere come agire per contrastare l'insorgenza della resistenza.

In questa ricerca, verranno analizzati tutti gli esperimenti effettuati e selezionati i dati su cui condurre il fitting, per ricavare i parametri migliori da inserire nel modello. Il modello verrà successivamente costruito sia in assenza che in presenza di antibiotico e verranno eseguite simulazioni a tempi successivi a quelli sperimentali, variando i parametri legati all'effetto dell'antibiotico e alla coniugazione del plasmide.

3 Materiali e Metodi

3.1 R

R è un linguaggio di programmazione utilizzato in ambito scientifico per le analisi di tipo statistico e la manipolazione e rappresentazione grafica dei dati. In questa ricerca, è stato utilizzato per analizzare e rappresentare graficamente i dati sperimentali, ricavando informazioni sulla dinamica delle comunità e sui tassi di crescita dei singoli ceppi all'interno di diversi terreni di coltura.

RStudio è un ambiente di sviluppo integrato (IDE) per R ed è il software utilizzato per condurre le analisi e le rappresentazioni grafiche dei dati, nella versione 2026.01.0.

Le librerie di R utilizzate sono necessarie per organizzare e rappresentare graficamente i dati e sono le seguenti:

1. *ggplot2* [9]: Questa libreria fa parte dell'ecosistema *Tidyverse* ed è utilizzata per rappresentare graficamente i dati. Permette di visualizzare grafici complessi, modificare esteticamente lo stile grafico e aggiungere a strati tutti gli elementi desiderati, che possono anche sovrapporsi.
2. *ggpubr* [10]: Questa libreria è utilizzata in concomitanza con il pacchetto *ggplot2* per la rappresentazione grafica dei dati. Permette di semplificare la sintassi del pacchetto *ggplot2* e di combinare insieme più grafici all'interno di un'unica immagine attraverso la funzione *ggarrange*.
3. *dplyr* [11]: Questa libreria fa parte dell'ecosistema *Tidyverse* ed è utilizzata per la manipolazione dei dati. Permette di ridurre un insieme di osservazioni a un singolo valore di sintesi tramite la funzione *summarize*.
4. *pals* [12]: Questa libreria è utilizzata per aggiungere palette di colori ai grafici *ggplot* e ottenere gli stessi colori presenti nei grafici di MATLAB.

3.2 MATLAB

MATLAB è una piattaforma di programmazione e calcolo per l'analisi e la visualizzazione dei dati e per l'esecuzione di tante operazioni matematiche, tra cui le regressioni sui dati. In questa ricerca, è stato utilizzato per costruire un modello matematico batterico funzionante composto da equazioni differenziali. Inoltre, è stato usato per manipolare i dati, così da filtrare e analizzare soltanto quelli di interesse, e per svolgere operazioni matematiche più complesse, tra cui le regressioni esponenziali e il fitting sui dati sperimentali. La versione di MATLAB utilizzata è MATLAB R2025b.

I toolbox di MATLAB utilizzati aggiungono funzioni matematiche ulteriori a quelle presenti nel pacchetto base di MATLAB e permettono di ottenere le regressioni esponenziali e calcolare gli intervalli di confidenza al 95% dei parametri ottenuti dal fitting sui dati:

1. *Curve Fitting Toolbox* [13]: Questo toolbox è utilizzato per ottenere le regressioni matematiche attraverso la funzione *curveFitter*, che apre un'interfaccia grafica.
2. *Statistics and Machine Learning Toolbox* [14]: Questo toolbox è usato per calcolare gli intervalli di confidenza al 95% dei parametri ottenuti attraverso il fitting tramite la funzione *nlparci*.

3.2.1 Fitting sui dati sperimentali

Il modello, che rappresenta la comunità formata da tre ceppi batterici, è composto da sei equazioni differenziali, che contengono al loro interno alcuni parametri, come i tassi di crescita e di morte dei singoli ceppi batterici e i tassi di produzione degli aminoacidi. Poiché questi parametri descrivono grandezze biologiche misurabili, possono essere stimati a partire dai dati sperimentali.

Per ottenere un modello che rappresenti al meglio i dati sperimentali catturandone l'andamento, è necessario eseguire due passaggi. Nel primo, si utilizza una funzione di fitting per calcolare i parametri ottimizzati del modello, ossia quelli che approssimano meglio i dati sperimentali. Nel secondo, si utilizza la funzione *ode45* per risolvere numericamente il sistema di equazioni differenziali, utilizzando i parametri ottenuti nel primo passaggio e producendo simulazioni del modello confrontabili con i dati sperimentali.

SIMPLEXL

Per ottenere i parametri migliori da utilizzare nel fitting, esistono varie funzioni che cercano di ridurre il più possibile la distanza tra i dati sperimentali e i valori predetti dal fitting. La funzione di ottimizzazione utilizzata è *SIMPLEXL*, che è un metodo per trovare il minimo di una funzione a più variabili e utilizza il metodo del semplice con la fase di espansione ottimizzata. E' stata sviluppata da Aldo Buttini l'11 settembre 1993. La funzione *SIMPLEXL* richiede nella sintassi la funzione da ottimizzare, che in questo caso è *Objective Function* (paragrafo 3.2.1), i parametri di partenza (p_0), da cui poi calcolerà quelli migliori, le opzioni e i possibili parametri aggiuntivi.

La funzione costruisce un semplice iniziale, ossia una figura di $n+1$ vertici in uno spazio di n dimensioni, in cui gli $n+1$ vertici corrispondono all'insieme dei parametri iniziali p_0 diminuiti tutti del 20% e aumentati uno alla volta del 20%. Essendoci 7 parametri iniziali (tre tassi di crescita batterici, tre tassi di produzione degli aminoacidi e un tasso di morte), i vertici misurati nel semplice, ossia le combinazioni di parametri, sono otto e su ognuno viene calcolata la *objective function*. Successivamente, i vertici vengono ordinati dal migliore (*objective function* più bassa) al peggiore e l'algoritmo svolge più iterazioni in successione per sostituire il vertice peggiore con uno migliore. Via via che si sostituiscono i vertici con valori migliori, il semplice si avvicina al minimo. Inizialmente viene calcolato il baricentro di tutti i vertici tranne il peggiore, poi viene svolta una di queste possibili mosse:

1. Riflessione: il vertice peggiore viene riflesso rispetto al baricentro nella direzione opposta dello spazio. Se il nuovo punto è migliore del penultimo vertice, ma non del primo, viene accettato, mentre se è migliore del primo, avviene un'espansione.
2. Espansione: se la riflessione ha dato un buon risultato, allora il minimo potrebbe essere nella zona esplorata, quindi viene allungato il passo in quella direzione trovando

un altro punto. Se questo nuovo punto è ancora migliore, avviene un'espansione ottimizzata, ossia si continua ad andare in quella direzione con altri passi. Quando il punto ottenuto peggiora, l'algoritmo considera tutti i punti ottenuti nelle varie espansioni e cerca di restringere l'intervallo per ottenere un punto preciso in quella zona.

3. Contrazione: se la riflessione non ha dato un buon risultato, ossia il nuovo punto ha un risultato della *objective function* peggiore del penultimo punto, la direzione in cui l'algoritmo sta cercando si allontana dal minimo, quindi viene cercato un nuovo punto più vicino al baricentro dei vertici, verso l'interno del semplice, che viene stretto.
4. Riduzione: se la contrazione non ha portato un buon risultato, l'algoritmo sposta ogni vertice tranne il migliore a metà strada tra i punti iniziali e il punto migliore, così da ridurre la grandezza del semplice.

Queste mosse vengono ripetute finché non si raggiunge il numero massimo di iterazioni, indicato da `options(14)`, parametro della funzione *foptions* (paragrafo 3.2.1), oppure fino a quando si raggiunge la tolleranza sui parametri o sul valore della funzione obiettivo, che sono indicati dalle `options(2)` e `options(3)` della funzione *foptions*. Il raggiungimento della tolleranza indica l'arrivo a un minimo locale, poiché tutte le funzioni obiettivo dei vertici hanno valori simili tra di loro. Il risultato che si ottiene dalla funzione *SIMPLEXL* è l'insieme dei parametri ottimizzati.

Objective Function

La funzione obiettivo rappresenta lo scarto che si trova tra il fitting ottenuto dal modello matematico e i dati sperimentali. E' la funzione che viene minimizzata da *SIMPLEXL* (paragrafo 3.2.1) e che riceve come input i dati sperimentali e i parametri da ottimizzare, ricavando le condizioni iniziali del sistema dai valori sperimentali al tempo zero. La funzione integra attraverso il solver *ode45* [15] le equazioni differenziali del modello e genera un grafico in tempo reale che mostra l'andamento del modello rispetto ai dati sperimentali. Le soluzioni ottenute dal modello vengono utilizzate per calcolare l'errore quadratico medio, ossia lo scarto tra il modello e i dati sperimentali per ogni ceppo batterico, mostrato dalle equazioni 3.1, 3.2 e 3.3, in cui err_A , err_B ed err_C sono i vettori degli errori quadratici punto per punto dei tre ceppi batterici, $solCfuA$, $solCfuB$ e $solCfuC$ sono le soluzioni ottenute dal modello e cfu_A , cfu_B e cfu_C sono i dati sperimentali, ossia le colony forming units (CFU).

$$err_A = (solCfuA - cfu_A)^2 \quad (3.1)$$

$$err_B = (solCfuB - cfu_B)^2 \quad (3.2)$$

$$err_C = (solCfuC - cfu_C)^2 \quad (3.3)$$

A partire dagli errori quadratici medi è possibile calcolare lo scarto attraverso l'equazione 3.4, in cui $length(time)$ è il numero di istanti di tempo in cui sono stati misurati i dati sperimentali e serve a normalizzare lo scarto, rendendolo indipendente dal numero di dati sperimentali. Questo risultato è il valore numerico che la funzione *SIMPLEXL* cerca di minimizzare modificando i parametri.

$$scarto = \frac{1}{length(time)} * \sum_{i=1}^N (err_A + err_B + err_C) \quad (3.4)$$

Dalla funzione obiettivo è possibile ricavare anche il vettore dei residui, che contiene tutte le differenze tra le soluzioni del modello e i dati sperimentali e che è utilizzato dalla funzione *nlparci* (paragrafo 3.2.2) per ottenere l'intervallo di confidenza al 95% di ognuno dei parametri.

Foptions

La funzione *foptions* comunica alla funzione *SIMPLEXL* (paragrafo 3.2.1) le opzioni da utilizzare per il suo funzionamento, come le tolleranze sui valori o il numero di valutazioni. Consiste in 18 opzioni che possono essere modificate dall'utente. Se l'utente non modifica nessun valore, vengono prese le 18 opzioni di default, mentre se inserisce solo alcune opzioni, *foptions* utilizza quei valori e usa i parametri di default per quelli mancanti. Delle 18 opzioni, quelle utilizzate dalla funzione *SIMPLEXL* sono le seguenti:

1. options(1): questa opzione permette di visualizzare i risultati intermedi durante l'esecuzione della funzione. Il valore di default è zero e non mostra i risultati, mentre cambiando questo valore vengono visualizzati.
2. options(2): questa opzione è il valore di tolleranza sui parametri calcolati, che di default vale 10^{-4} .
3. options(3): questa opzione è il valore di tolleranza sulla funzione utilizzata, ossia sulla *objective function* (paragrafo 3.2.1). Di default è 10^{-4} .
4. options(6): questa opzione stabilisce come viene inizializzato il semplice di partenza. Se vale 0, come avviene di default, il semplice viene sempre ricostruito a partire dal punto iniziale; se vale 1 o 3, viene invece recuperato da un file esterno (*STARTV*) contenente un semplice salvato in precedenza. Nel fitting eseguito in questa ricerca, è stato mantenuto il valore di default, per cui il semplice viene sempre calcolato da zero a partire da p0.
5. options(14): questa opzione è il numero massimo di iterazioni che svolge *SIMPLEXL* prima di terminare. Il valore di default di *foptions* vale zero, tuttavia, quando *SIMPLEXL* riceve questo valore nullo, lo sostituisce con 200 volte il numero di parametri da ottimizzare, che rappresenta quindi il numero massimo di iterazioni effettivamente utilizzato di default.

Ulteriori metodi per ottimizzare i parametri

Esistono ulteriori possibili metodi per ottenere i parametri ottimizzati a partire da dei valori iniziali. Esempi di questi metodi sono *SIMPSEA*, *fmincon* e *fminsearch*, tuttavia dopo averli testati è apparso evidente che i parametri migliori si ottenevano tramite la funzione *SIMPLEXL* (paragrafo 3.2.1), per cui non sono stati utilizzati. Questa scelta è dovuta sia ai tempi minori di esecuzione di *SIMPLEXL* rispetto agli altri metodi che ai fitting ottenuti con i parametri migliori calcolati, che rappresentavano in modo più veritiero l'andamento dei dati sperimentali.

ode45

Una volta ottenuti i parametri migliori attraverso la funzione di ottimizzazione *SIMPLEXL* (paragrafo 3.2.1), è possibile utilizzarli per risolvere numericamente il sistema di equazioni differenziali del modello e ottenere le curve simulate da confrontare con l'andamento dei dati sperimentali. La funzione utilizzata a questo scopo è *ode45* [15], che permette di risolvere le equazioni differenziali non rigide. Nella sua sintassi, richiede la funzione con le equazioni differenziali del modello, che include a sua volta i parametri migliori ottenuti dalla funzione di ottimizzazione, l'intervallo di tempo da simulare, le condizioni iniziali del sistema, ossia i dati sperimentali al tempo zero, ed eventuali opzioni aggiuntive. Attraverso *ode45* vengono calcolati i punti predetti dal modello nel tempo ed è possibile confrontarli graficamente con i punti sperimentali, così da valutare se il modello descrive realmente l'andamento dei dati sperimentali. Nel caso in cui il risultato non sia soddisfacente, è necessario modificare i parametri iniziali (p_0) dati alla funzione *SIMPLEXL* e ricominciare l'ottimizzazione, trovando nuovi parametri ottimizzati e utilizzandoli nuovamente all'interno della funzione *ode45* per valutare il nuovo risultato ottenuto. La funzione *ode45* permette di scegliere qualsiasi intervallo temporale, per cui è possibile simulare l'andamento del modello a tempi più lunghi rispetto a quelli sperimentali. Questo può essere utile per ricavare informazioni nel lungo periodo sulla comunità batterica studiata, predicendo ad esempio il raggiungimento di uno stato stazionario in cui i ceppi batterici convivono senza modifiche sostanziali delle loro concentrazioni, la morte cellulare dei ceppi oppure la loro crescita esponenziale.

3.2.2 Intervalli di confidenza

Gli intervalli di confidenza al 95% indicano la precisione della stima di ogni parametro ottenuto attraverso la funzione di ottimizzazione *SIMPLEXL* (paragrafo 3.2.1). Se l'intervallo è stretto, il parametro è affidabile, mentre se è largo, lo stesso fit può essere ottenuto con valori differenti del parametro, per cui la stima è meno affidabile. Considerare l'affidabilità dei parametri è importante perché, se sono affidabili, il modello descrive con precisione adeguata i dati sperimentali.

in italiano

Per stimare gli intervalli di confidenza, Statistics and Machine Learning Toolbox offre la funzione *nlparci* [16], che sta per **nonlinear regression parameter confidence intervals** e che stima la covarianza dei parametri linearizzando il modello intorno alla soluzione ottimale. Nella sua sintassi, richiede le stime dei parametri ottimizzati ottenuti tramite *SIMPLEXL*, il vettore dei residui, cioè le differenze tra i dati sperimentali e i valori predetti dal modello con i parametri stimati, e lo Jacobiano, una matrice che indica quanto cambia ogni residuo se si modifica leggermente ogni parametro. *nlparci* costruisce la matrice di covarianza dei parametri, considerando quanto il modello si discosta dai dati, ossia il vettore dei residui, e la sensibilità del modello a ogni parametro, vista dallo Jacobiano, e successivamente stima gli intervalli di confidenza.

3.2.3 Regressione

La regressione matematica è un metodo per stimare come cambia una variabile dipendente (y) al variare di una variabile indipendente (x), ottenendo una formula matematica che permette di descrivere la relazione tra le due variabili e di fare previsioni [17]. La curva ottenuta tramite la regressione è quella che più si avvicina a tutti i dati e attraverso la sua equazione è possibile stimare i nuovi valori di y conoscendo i relativi valori di x .

Esistono vari tipi di regressione, come lineare, esponenziale, logaritmica, polinomiale, sigmoideale e gaussiana, e ognuna descrive un diverso andamento della relazione tra le variabili x e y . Quella utilizzata in questa ricerca è la regressione esponenziale.

Su MATLAB, è possibile ottenere le regressioni attraverso la funzione *curveFitter* [13], che apre un'interfaccia grafica e permette di selezionare i valori di x e y desiderati. Dopo averli rappresentati in un piano cartesiano, è possibile scegliere tra più tipi di regressione, selezionando quello che meglio rappresenta la relazione tra i dati. I risultati ottenuti sono la funzione che descrive la regressione, i coefficienti da inserire all'interno della funzione con il loro relativo intervallo di confidenza al 95% e diversi parametri correlati alla qualità del fit, come SSE, R-square, DFE, Adjusted R-square e RMSE. Questi cinque valori sono parametri statistici che indicano quanto un modello matematico rappresenta l'andamento dei dati e che misurano la distanza che c'è tra i punti sperimentali e quelli predetti dal modello.

SSE (Sum of Squared Errors) misura la deviazione totale dei valori predetti dal fit rispetto a quelli sperimentali. Più si avvicina a zero, migliore è il fit, essendo bassa la componente di errore casuale. Si misura con l'equazione 3.5, in cui w_i è il peso associato a ogni deviazione, y_i è il valore osservato e $\hat{y}_i$ è il valore previsto.

$$SSE = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2 \quad (3.5)$$

R-square misura quanto il modello spiega la variazione dei dati e un fit migliore corrisponde a un R-square più alto. Per esempio, se R-square è 0.8, il fit spiega l'80% di tutta la variazione presente nei dati. Viene misurato attraverso l'equazione 3.6, sottraendo a 1 il rapporto tra la variabilità non spiegata dal modello (SSE) e la variabilità totale dei dati rispetto alla media (SST, equazione 3.7).

$$R^2 = 1 - \frac{SSE}{SST} \quad (3.6)$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (3.7)$$

DFE (Degrees of Freedom for Error) indica quanti gradi di libertà rimangono nei dati per stimare la variabilità residua, dopo che il modello ha utilizzato parte dei gradi di libertà per stimare i suoi m parametri. Si ottiene dalla formula 3.8 sottraendo il numero di coefficienti calcolati dal modello di regressione (m) al numero di dati (n).

$$\nu = n - m \quad (3.8)$$

Adjusted R-square è un parametro che permette di comparare modelli che hanno un diverso numero di variabili. Penalizza l'aggiunta di nuove variabili che non aiutano realmente il modello e di solito è minore di R-square. Si ottiene dall'equazione 3.9 e anche in questo caso, maggiore è il suo valore, migliore è il fit.

$$adjusted\ R^2 = 1 - \frac{SSE(n-1)}{SST \cdot \nu} \quad (3.9)$$

RMSE (Root Mean Squared Error) è una stima di quanto il modello si allontana nelle predizioni dai dati reali e si misura tramite l'equazione 3.10. Il fit migliore ha un basso RMSE [18].

$$RMSE = \sqrt{\frac{SSE}{\nu}} \quad (3.10)$$

3.3 Dati sperimentali

I dati sperimentali sono stati forniti dal gruppo di ricerca di Microbial Pangenomes, guidato dal professor Marco Galardini, con sede presso il TWINCORE (Centro per la ricerca sperimentale e clinica sulle infezioni) ad Hannover. Gli strumenti utilizzati per ottenere i dati sperimentali sono il Chi.Bio (paragrafo 3.3.1), il plate counting (paragrafo 3.3.2) e il citofluorimetro (paragrafo 3.3.3); inoltre, sono state utilizzate piastre da 96 pozzetti per ricavare il tasso di crescita dei singoli ceppi in mezzi differenti.

3.3.1 Chi.Bio

Le comunità batteriche sono state cresciute all'interno dello strumento Chi.Bio, una piattaforma sperimentale open-source utilizzata per automatizzare gli esperimenti di biologia sintetica, dei sistemi ed evolutiva. Chi.Bio permette di monitorare e mantenere costanti i parametri dell'ambiente in cui crescono i batteri, aumentando la riproducibilità degli esperimenti, che altrimenti risentirebbe della variazione di condizioni ambientali come la temperatura e la velocità di agitazione delle colture cellulari [19].

I dati che vengono misurati automaticamente da questo strumento sono i seguenti:

1. Densità ottica (OD), che indica la quantità di luce assorbita da una soluzione e che è legata al numero di cellule presenti nella soluzione misurata. Una maggiore densità ottica indica una maggiore presenza cellulare nella soluzione, anche se non permette di distinguere le cellule in vita da quelle morte.
2. Fluorescenza (FP1 e FP2). In questi esperimenti vengono monitorati due segnali di fluorescenza, appartenenti a due dei tre ceppi delle comunità batteriche.
3. Temperatura della soluzione, che viene costantemente monitorata per mantenerla intorno alla temperatura ottimale di crescita batterica, ossia 37 °C, evitando riscaldamenti e raffreddamenti.
4. Velocità delle pompe in ingresso e in uscita del Chi.Bio. La pompa in ingresso introduce il terreno fresco e pieno di nutrienti all'interno dello strumento, mentre la pompa in uscita espelle il terreno di scarto. Queste pompe sono necessarie per mantenere le cellule sempre nelle condizioni ottimali per la loro crescita, in modo che le oscillazioni di crescita osservate nel tempo siano dovute solamente alle relazioni presenti nelle comunità batteriche e non a fattori esterni.
5. Velocità di mescolamento delle colture cellulari, che deve essere mantenuta costante per evitare una diminuzione della loro ossigenazione.

3.3.2 Plate counting

Il plate counting è uno strumento che permette di contare il numero di CFU (colony forming units), ossia il numero di colonie batteriche presenti in una piastra. Attraverso la fluorescenza dei ceppi PM (rossa) e PL (verde) è possibile contare le CFU di ogni ceppo singolarmente, ottenendo informazioni sulla loro crescita.

Per ottenere i dati sulle CFU, è stata estratta parte della soluzione cellulare interna al Chi.Bio ed è stata piastrata, ripetendo questa operazione circa una volta al giorno. Dopo un periodo di incubazione, ogni piastra è stata letta dal plate counting, ottenendo i dati sperimentali.

3.3.3 Citometria a flusso

Il citofluorimetro è uno strumento che permette di misurare il numero di eventi, ossia di cellule, direttamente in soluzione e senza necessità di piastrare e incubare le cellule. La soluzione estratta dal Chi.Bio viene fatta passare direttamente all'interno dello strumento, che misura il numero totale di eventi e il numero di eventi per singolo ceppo, ricavando anche la proporzione di ogni ceppo rispetto agli eventi totali.

3.4 Comunità microbica sintetica

Se vogliamo lasciare il termine inglese dobbiamo mettere knocked out, in italico. Altrimenti possiamo dire che “è deletato in due geni”

La comunità microbica studiata è una comunità sintetica stabile composta da tre ceppi di *Escherichia coli* BW25113. Ognuno dei tre ceppi è **knockout** per due geni necessari alla biosintesi di aminoacidi e riceve gli aminoacidi che non è in grado di produrre, dagli altri due ceppi (Figura 3.1). La comunità è costituita in modo tale che i tre ceppi stabiliscano relazioni mutualistiche tra loro e possano sopravvivere soltanto se sono tutti presenti. In mancanza di uno dei ceppi, non viene prodotto uno degli aminoacidi fondamentali per gli altri due, che quindi non riescono a sopravvivere da soli. I tre ceppi studiati sono LM (Δ **lysA**, Δ metA), che non produce lisina e metionina, PM (Δ pheA, Δ metA), che non produce fenilalanina e metionina, e PL (Δ pheA, Δ lysA), che non produce fenilalanina e lisina. La scelta dei tre ceppi si basa sullo studio di Mee et al. [4], che identifica questa combinazione di ceppi auxotrofi come una delle combinazioni a maggiore resa di crescita in terreni minimi senza aminoacidi. I ceppi PM e PL contengono due plasmidi piccoli e non mobili che esprimono due marcatori di fluorescenza: mScarlet nel caso di PM, che esprime una fluorescenza rossa, e GFP nel caso di PL, che esprime una fluorescenza verde. Attraverso le due fluorescenze è possibile distinguere i tre ceppi all'interno della coltura totale.

I tre ceppi sono stati coltivati in terreno M9 all'interno dello strumento Chi.Bio in colture continue e sono stati monitorati costantemente tramite sensori di densità ottica e fluorescenza per seguirne la crescita e l'andamento dinamico. M9 è un terreno minimo, composto da fosfati di sodio e potassio, cloruro di sodio e cloruro di ammonio, in cui i membri della comunità devono instaurare relazioni per produrre tutti gli aminoacidi, che non sono presenti nel terreno, e garantirsi la sopravvivenza. Tutti i ceppi presentano resistenza alla kanamicina, che è un antibiotico utilizzato per ottenere la crescita selettiva dei ceppi sulle piastre ed evitare la contaminazione da parte di altre specie non resistenti. La kanamicina è un antibiotico che appartiene alla classe degli aminoglicosidi ed è attivo soprattutto nei confronti dei Gram negativi. Agisce legando la subunità ribosomiale 30S e bloccando la sintesi proteica, che causa la morte del batterio [20].

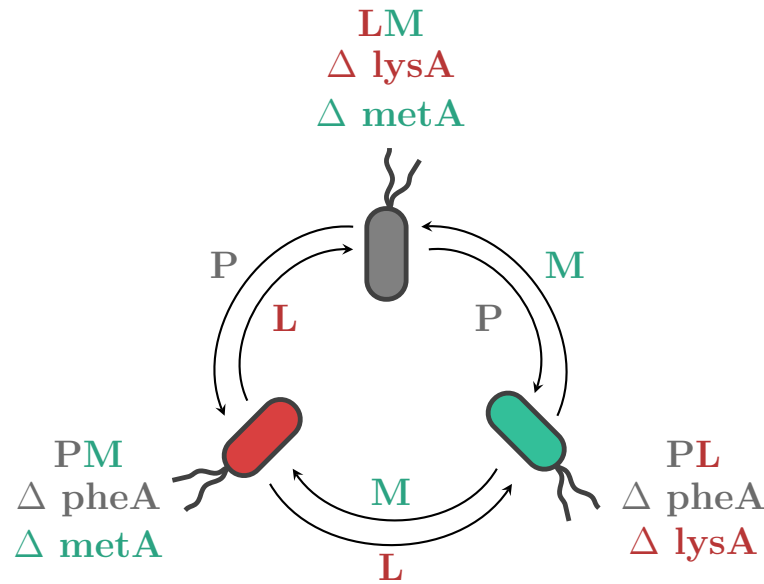


Figura 3.1: Schema delle interazioni mutualistiche presenti tra i ceppi LM, PM e PL, che permettono la sopravvivenza della comunità solo in presenza di tutti i ceppi.

3.4.1 Antibiotici utilizzati

Gli antibiotici utilizzati per studiare l'antibiotico-resistenza sono ampicillina e colistina. L'ampicillina appartiene alla classe dei β -lattamici, in particolare alla famiglia delle penicilline, ed è un antibiotico ad ampio spettro utilizzato per trattare infezioni causate da batteri Gram positivi e alcuni Gram negativi, come *Escherichia coli*. Agisce sulla sintesi del peptidoglicano, che è la componente principale della parete cellulare batterica, andando a legare le penicillin-binding proteins (PBP). Questi enzimi catalizzano la transpeptidazione, ossia la formazione di cross-link tra le catene di peptidoglicano, e legano in modo irreversibile l'anello β -lattamico dell'ampicillina, che è un mimetico del substrato naturale di questi enzimi, ossia del dipeptide terminale D-alanil-D-alanina [21]. Inattivando le PBP, la parete cellulare non è più stabile e i batteri muoiono per lisi osmotica.

La colistina (polimixina E), invece, è un antibiotico appartenente alla classe delle polimixine ed è utilizzato contro i batteri Gram negativi. In terapia, il suo utilizzo è limitato ai casi di batteri Gram negativi multiresistenti, a causa della forte tossicità che comporta. Agisce sul lipide A del lipopolisaccaride batterico, una componente fondamentale della membrana esterna dei Gram negativi, legandosi in modo elettrostatico e destabilizzando la struttura della membrana esterna fino a causare la morte del batterio [22].

I plasmidi che causano resistenza ai due antibiotici sono pOXA48 e PN23. pOXA48 codifica per la carbapenemasi OXA-48, che idrolizza l'anello β -lattamico dell'ampicillina e inattiva l'antibiotico [23]; PN23, invece, è un plasmide a incompatibilità IncX4 che porta il gene *mcr-1*, codificante per una fosfoetanolamina transferasi che modifica il lipide A del lipopolisaccaride, riducendo l'affinità della colistina per il suo bersaglio [24].

3.4.2 Tipi di comunità per lo studio di antibiotico-resistenza con ampicillina e colistina

I tre ceppi sono stati combinati a formare diverse comunità per studiare il meccanismo di antibiotico-resistenza attraverso il trasferimento genico orizzontale. Nello studio della resistenza all'ampicillina, sono state considerate quattro comunità (Figura 3.2). La prima, chiamata KAN, contiene tutti i ceppi sensibili all'antibiotico e permette di studiare l'andamento di crescita della colonia batterica in mancanza di plasmidi di resistenza, che possono avere un impatto sulla fitness. Rappresenta il controllo positivo dell'esperimento. La seconda, chiamata KAN + PVI, è formata inizialmente da tutti i ceppi sensibili, ma dopo due ore viene aggiunto il ceppo LM contenente la variante PVI del plasmide pOXA48, che gli conferisce la resistenza e che può essere trasferito ai ceppi sensibili [23]. Attraverso questa comunità vengono studiati l'impatto dell'antibiotico in assenza di resistenza e il cambiamento della dinamica di comunità successivo all'aggiunta di un ceppo resistente capace di trasferire la resistenza. Trasferendo il plasmide, il ceppo resistente permette, tramite coniugazione, l'acquisizione del plasmide e la sopravvivenza dei ceppi sensibili, che portano alla sopravvivenza dell'intera comunità. La terza comunità, chiamata PVI, contiene il ceppo LM resistente grazie alla presenza della variante PVI del plasmide pOXA48 e i ceppi PM e PL sensibili. LM può trasferire il plasmide e rendere gli altri due ceppi resistenti. L'ultima comunità, chiamata PVB, ha il ceppo LM resistente contenente la variante PVB del plasmide pOXA48, che non può essere coniugato a causa della mancanza dei geni per la coniugazione [23], mentre i ceppi PM e PL sono sensibili. In questo caso, LM non può trasmettere la resistenza agli altri ceppi, che quindi risultano sensibili all'effetto dell'antibiotico. Questa comunità rappresenta il controllo negativo e permette di osservare l'antibiotico-resistenza non genica. Poiché le carbapenemasi prodotte da pOXA48 degradano l'ampicillina e offrono un effetto protettivo sui ceppi sensibili, portando a una resistenza non genica, attraverso il confronto tra la comunità con il plasmide PVB e le altre è possibile osservare se il trasferimento genico orizzontale è trascurabile rispetto all'effetto protettivo dovuto alle carbapenemasi oppure se è il fattore principale che causa l'antibiotico-resistenza. Ognuna delle quattro comunità descritte è stata studiata in presenza e in assenza di antibiotico per osservare come la presenza e l'assenza dell'antibiotico e dei plasmidi impattano sulla fitness dei ceppi e influenzano la comunità e la sua composizione.

Cosa è la
antibiotico-
resistenza non
genica?
spiegare meglio
o rimuovere

Nello studio della resistenza alla colistina, sono state considerate tre comunità (Figura 3.3). La prima è KAN e corrisponde a quella studiata per la resistenza all'ampicillina, poiché tutti i ceppi sono sensibili. La seconda è KAN + PN23, che è il corrispettivo di KAN + PVI e inizialmente è formata da tutti i ceppi sensibili, mentre dopo due ore viene aggiunto il ceppo LM contenente il plasmide PN23, che gli conferisce la resistenza. LM può trasferire il plasmide agli altri due ceppi e rendere tutta la comunità resistente, in modo da farla sopravvivere all'antibiotico. La terza comunità è PN23, che è analoga a PVI e contiene il ceppo LM resistente per la presenza del plasmide PN23 e i ceppi PM e PL sensibili. Anche in questo caso, il plasmide PN23 può essere trasferito per coniugazione ai ceppi sensibili. Non è presente, invece, la quarta comunità in cui non può avvenire il trasferimento del plasmide poiché la resistenza alla colistina si basa su un meccanismo diverso dal rilascio di enzimi extracellulari, quindi non è necessario controllare i meccanismi di resistenza basati su un effetto protettivo.

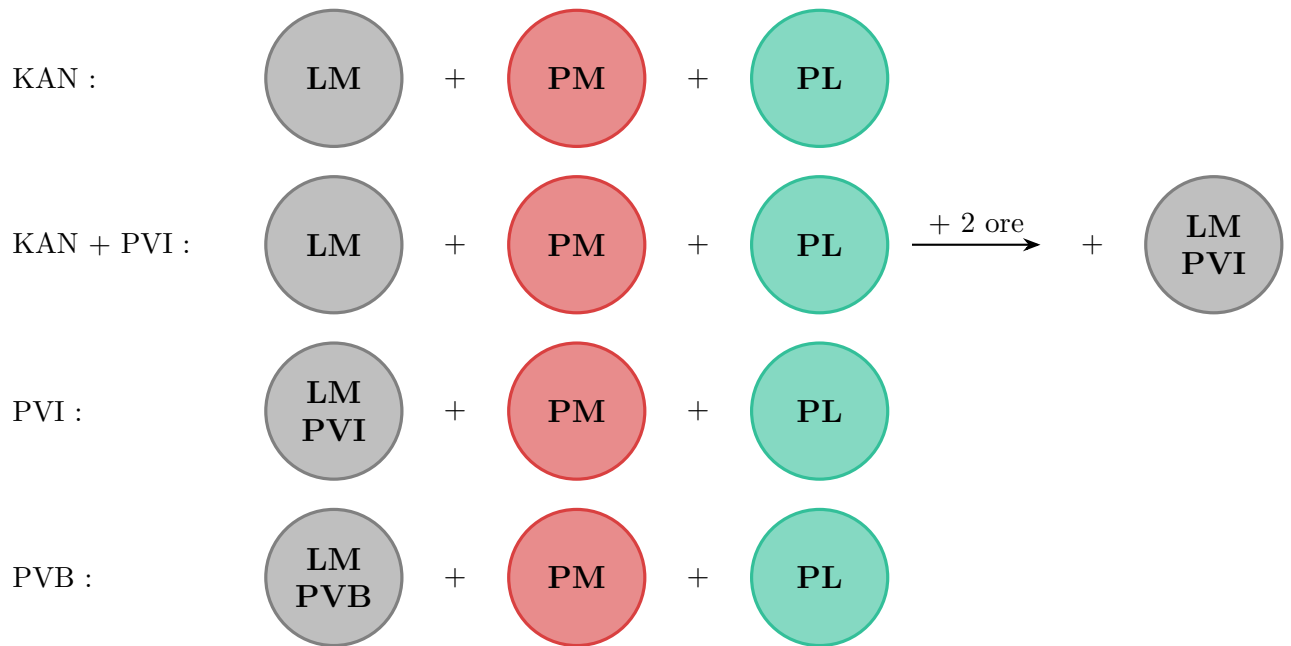


Figura 3.2: Rappresentazione delle quattro comunità con i relativi ceppi per lo studio della resistenza all'ampicillina.

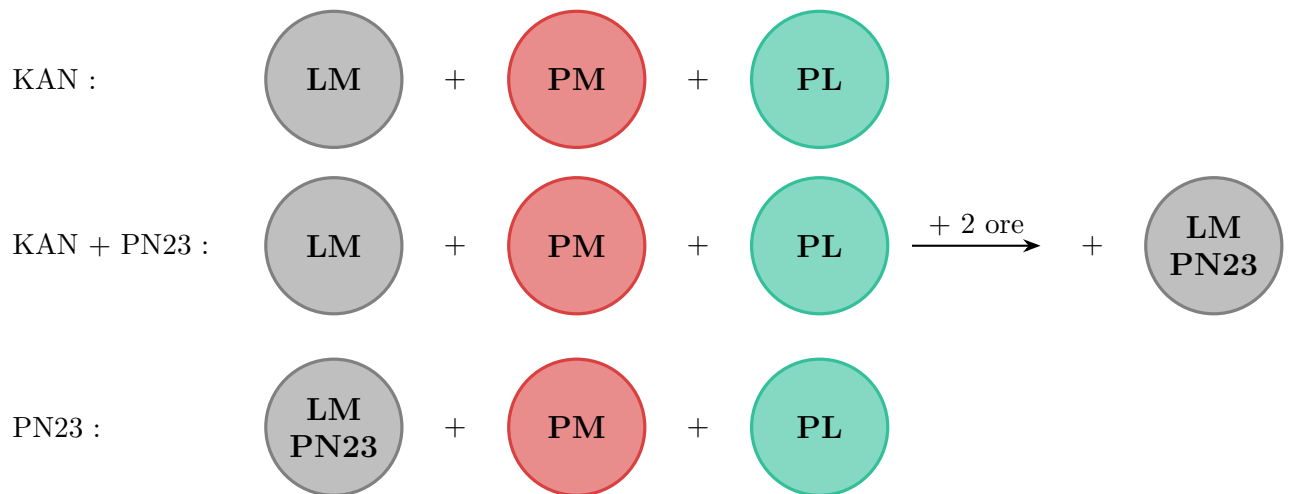


Figura 3.3: Rappresentazione delle tre comunità con i relativi ceppi per lo studio della resistenza alla colistina.

4 Risultati e discussione

4.1 Descrizione dei quattro esperimenti

Gli esperimenti eseguiti per studiare le comunità in assenza e in presenza di antibiotico sono quattro (Figura 4.1): uno è stato effettuato sulle comunità con i plasmidi di resistenza all'ampicillina, uno sulla comunità KAN e due sulle comunità con il plasmide di resistenza alla colistina.

Il primo esperimento è stato svolto sulle comunità KAN, KAN + PVI, PVI e PVB in presenza e in assenza di ampicillina, per un totale di otto colture, ed è composto da due replicati di durata temporale diversa: il primo replicato dura 208 ore, mentre il secondo 306. Il primo replicato della comunità KAN senza antibiotico ha una durata di sole 44 ore a causa di problemi sperimentali, che hanno portato all'interruzione della coltura. A causa della differenza di durata tra il primo e il secondo replicato e dei problemi che hanno portato all'interruzione del primo replicato della comunità KAN senza antibiotico, è stato analizzato soltanto il secondo replicato delle quattro comunità. L'ampicillina è stata aggiunta in concentrazione 16xMIC, utilizzando una concentrazione minima inibitoria (MIC) di 4 µg/ml. I dati sperimentali che sono stati ottenuti in questo esperimento sono la densità ottica e la fluorescenza tramite il Chi.Bio (paragrafo 3.3.1), le CFU tramite plate counting (paragrafo 3.3.2), e gli eventi tramite la citometria a flusso (paragrafo 3.3.3).

Il secondo esperimento è stato svolto soltanto sulla comunità KAN senza antibiotico e con un unico replicato di 498 ore. I dati sperimentali ricavati sono la densità ottica e la fluorescenza tramite il Chi.Bio.

Il terzo esperimento è stato eseguito sulle comunità KAN, KAN + PN23 e PN23: KAN e PN23 sono state coltivate in assenza e in presenza di colistina, testando concentrazioni di 1xMIC e 2xMIC, che corrispondono rispettivamente a 0.5 µg/ml e 1 µg/ml; KAN + PN23, invece, è stata coltivata solamente in presenza di 2xMIC di colistina. Sono presenti due replicati di 309 ore per ogni comunità, tranne che i secondi replicati di KAN, KAN + PN23 e PN23 in presenza di 2xMIC di colistina, che hanno una durata di 120 ore. I dati sperimentali ottenuti per entrambi i replicati sono le CFU tramite plate counting, mentre soltanto per il primo replicato sono presenti anche le misure di densità ottica e fluorescenza tramite Chi.Bio e degli eventi tramite la citometria a flusso.

L'ultimo esperimento è stato svolto sulle comunità KAN e PN23 in presenza e in assenza di colistina, utilizzata in concentrazioni 1xMIC e 2xMIC. Si trova un unico replicato con durata temporale diversa per le varie comunità: per le comunità KAN e PN23 in assenza di colistina, la durata è di circa 1000 ore; per le comunità KAN e PN23 con 1xMIC di colistina la durata è di 455 ore; la comunità KAN con 2xMIC di colistina ha durata di 140 ore; la comunità PN23 con 2xMIC di colistina ha durata di 260 ore. I dati ottenuti sono la densità ottica e la fluorescenza tramite Chi.Bio, le CFU tramite plate counting e gli eventi tramite citometria a flusso.

① → 306 h <i>pOXA48</i> Rep1 X Rep2 V Antibiotico: ampicillina Comunità: KAN no drug KAN 16xMIC KAN + PVI no drug KAN + PVI 16xMIC PVI no drug PVI 16xMIC PVB no drug PVB 16xMIC Dati: Densità ottica Fluorescenza Conte cellulari Eventi	② → 498 h Rep1 Comunità: KAN no drug Dati: Densità ottica Fluorescenza	③ → 309 h <i>PN23</i> Rep1 Rep2 (* = 120 h) Antibiotico: colistina Comunità: KAN no drug KAN 1xMIC KAN 2xMIC * KAN + PN23 2xMIC * PN23 no drug PN23 1xMIC PN23 2xMIC * Dati: Densità ottica (solo rep1) Fluorescenza (solo rep1) Conte cellulari Eventi (solo rep1)	④ → 1000 h <i>PN23</i> Rep1 Antibiotico: colistina Comunità: KAN no drug KAN 1xMIC (455 h) KAN 2xMIC (140 h) PN23 no drug PN23 1xMIC (455 h) PN23 2xMIC (260 h) Dati: Densità ottica Fluorescenza Conte cellulari Eventi
--	---	--	---

Figura 4.1: Rappresentazione dei quattro esperimenti con i numeri dei replicati, l'antibiotico utilizzato, le comunità esposte e i dati ottenuti nell'esperimento.

4.2 Analisi dati gli esperimenti

Sono stati analizzati i dati del primo esperimento, ossia densità ottica, fluorescenza dei ceppi PM e PL, CFU ed eventi, e sono stati visualizzati graficamente tramite l'utilizzo dei pacchetti *ggplot* e *ggpubr* di R (paragrafo 3.1). I dati di densità ottica, provenienti dal Chi.Bio, sono mostrati in figura 4.2 ed è possibile osservare che tutte le comunità in assenza di antibiotico (KAN, KAN + PVI, PVI e PVB) mostrano una dinamica oscillatoria nel tempo, aumentando e diminuendo le loro densità ottiche. Le comunità in presenza di antibiotico, invece, hanno andamenti diversi in base al tipo di comunità. La comunità KAN, avendo tutti ceppi sensibili, non sopravvive alla presenza di ampicillina, per cui va incontro alla morte cellulare con un andamento di decadimento. Le comunità KAN+PVI e PVI, invece, mostrano un andamento oscillatorio, dimostrando l'efficacia dell'antibiotico-resistenza indotta dal plasmide PVI nel ceppo LM e il suo avvenuto trasferimento ai ceppi PM e PL, che ha garantito la sopravvivenza di tutta la comunità. Infine, la comunità PVB mostra una crescita iniziale, dovuta alla presenza del ceppo LM resistente, ma non in grado di trasferire il plasmide, seguita da un decadimento a causa della morte dei ceppi PM e PL sensibili, che causa la fine anche del ceppo LM per la mancata produzione dei due aminoacidi forniti dai partner.

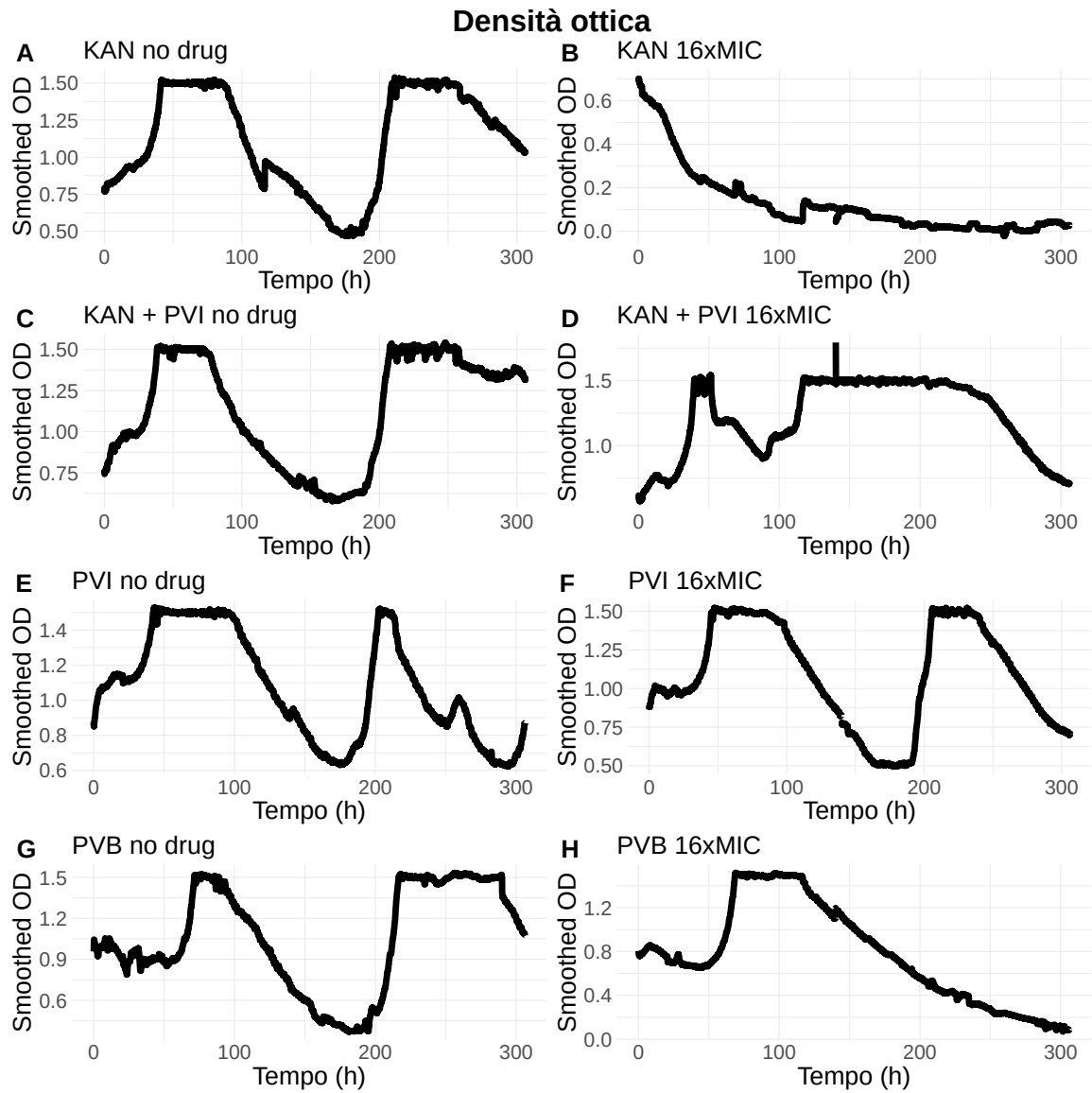


Figura 4.2: Densità ottica delle quattro comunità del primo esperimento in presenza e in assenza di ampicillina.

I dati di fluorescenza dei ceppi PM e PL, ottenuti dallo strumento Chi.Bio, sono rappresentati in figura 4.3 e ricalcano in parte l'andamento ottenuto dalla densità ottica. Le comunità in assenza di antibiotico mostrano andamenti oscillatori dei due ceppi PM e PL, che sopravvivono aumentando e diminuendo quasi sempre in fase. La comunità KAN in presenza di ampicillina non sopravvive, per cui le fluorescenze di PM e PL non presentano un andamento oscillatorio, ma di decadimento; le comunità KAN + PVI e PVI mostrano fluorescenze oscillatorie, per cui i loro ceppi sono vitali; la comunità PVB mostra un'oscillazione iniziale, seguita da un andamento piatto che corrisponde alla morte cellulare dei ceppi PM e PL.

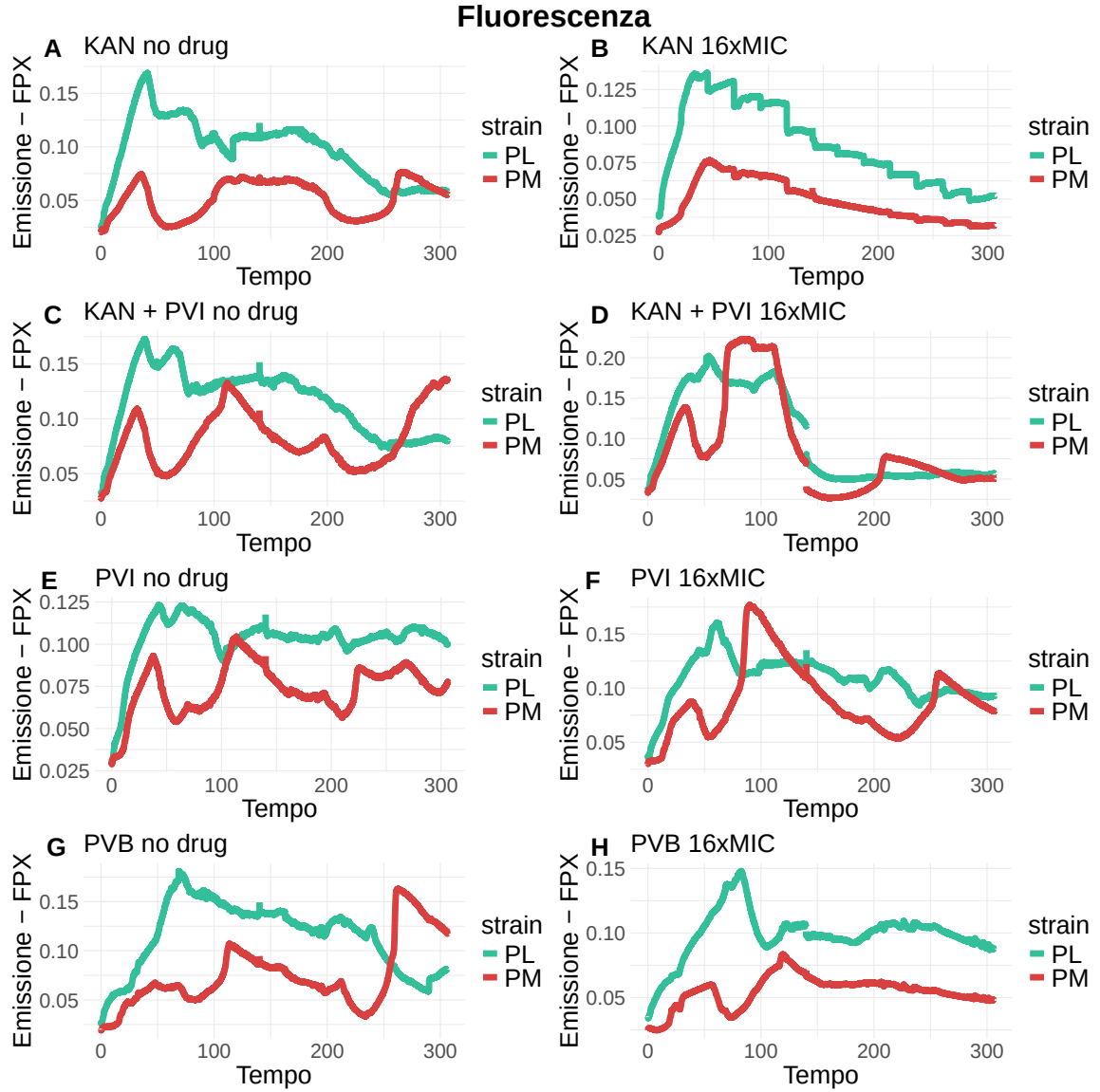


Figura 4.3: Fluorescenze delle quattro comunità del primo esperimento in presenza e in assenza di ampicillina.

Le CFU delle quattro comunità, mostrate in figura 4.4, rispecchiano i risultati ottenuti dai dati precedenti e osservano come cambia nel tempo l'abbondanza dei singoli ceppi. Le comunità in assenza di antibiotico mostrano tutte la presenza di due oscillazioni, la prima ottenuta poco prima di 100 ore e la seconda dopo 200 ore. La comunità KAN in presenza di ampicillina non sopravvive, presentando solo un piccolo picco del ceppo PM, seguito dalla morte cellulare di tutti i ceppi. Le comunità KAN + PVI e PVI hanno un andamento oscillatorio, dimostrando la sopravvivenza della comunità dovuta al passaggio del plasmide. La comunità PVB ha un picco iniziale di crescita, che è seguito dalla morte della comunità a causa dell'assenza del trasferimento del plasmide. Questa comunità mostra, sia per la densità ottica che per le conte cellulari, un andamento intermedio tra la comunità KAN e PVI: collassa a tempi successivi rispetto a KAN e precedenti rispetto a PVI, che rimane stabile, dimostrando che la degradazione dell'ampicillina da parte di LM rallenta l'effetto dell'antibiotico sui ceppi sensibili, ma non riesce a evitarne la morte. Dai dati di CFU è possibile osservare che il ceppo LM risulta sempre sottostimato, in

4 Risultati e discussione

Cosa si intende qui, le dimensioni delle cellule o che altro? Specificare

parte anche **a causa delle dimensioni minori del ceppo**. Sapendo che lo strumento di plate counting ha un limite di rilevazione dell'1%, tutti i ceppi in quantità inferiore all'1% rispetto al ceppo con il numero di cellule maggiore risultano avere CFU azzerate, come avviene per il ceppo LM durante la maggior parte dell'esperimento. Poiché i ceppi di PM e PL hanno un numero di cellule intorno a 10^8 , una popolazione di LM sotto a 10^6 cellule non viene rilevata, risultando nulla, e questo risulta problematico anche a livello concettuale, poiché PM e PL non possono sopravvivere in assenza di LM. Per superare questo limite, si impone una regola: se il numero di colonie è sotto una soglia di rilevazione, allora le CFU stimate sono la metà del limite di rilevazione. Attraverso questa strategia, le CFU di LM, indicate come CFU/mL Alt, non risultano più nulle, ma rimangono vicine allo zero in modo realistico.

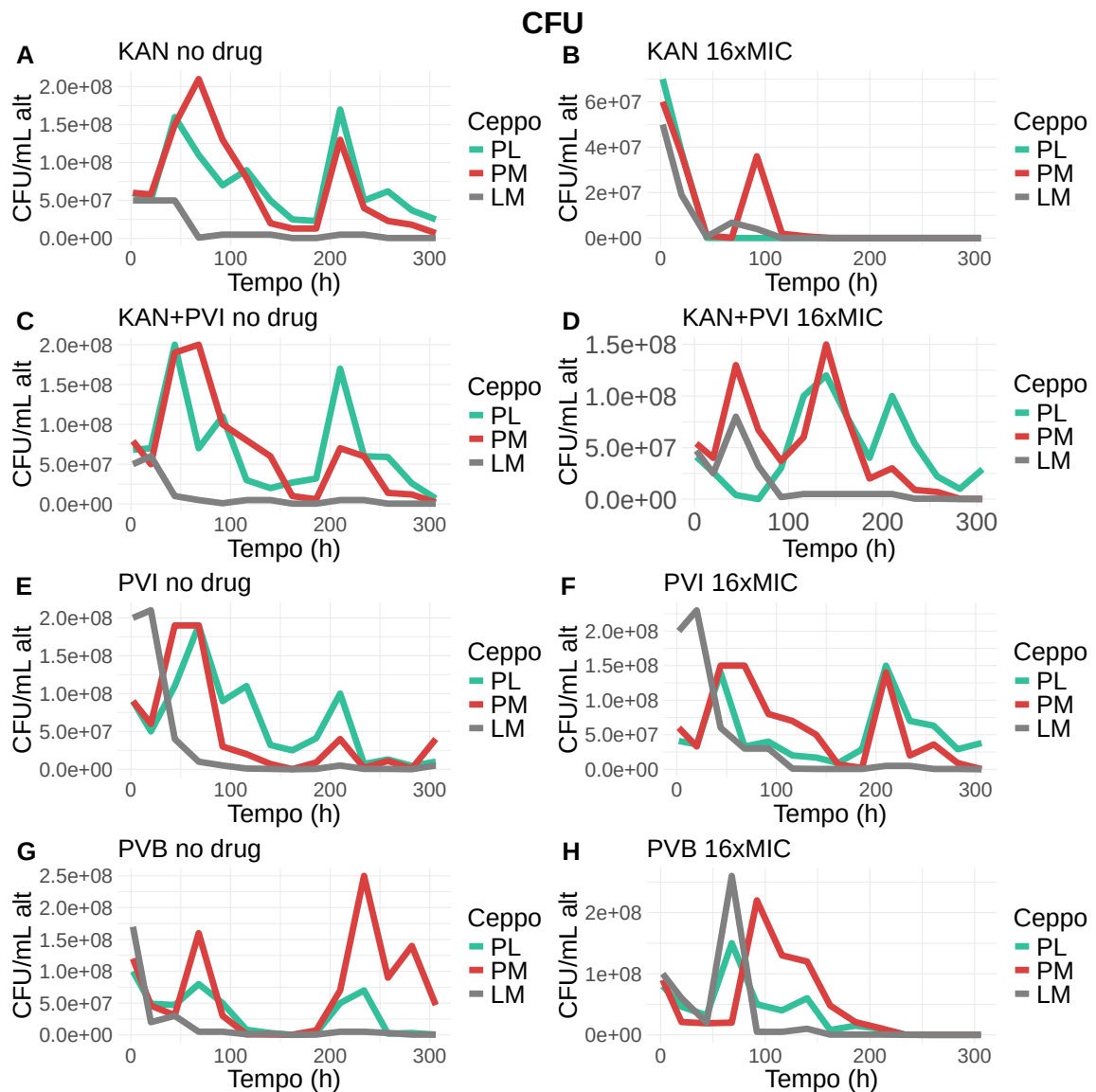


Figura 4.4: Plate counts delle quattro comunità del primo esperimento in presenza e in assenza di ampicillina.

I dati degli eventi ottenuti tramite la citometria a flusso sono mostrati in figura 4.5 e rispecchiano le oscillazioni e la morte cellulare presenti nei grafici delle CFU. E' possibile

notare che nella citometria a flusso la popolazione di LM è sovrastimata, contrariamente ai dati provenienti dal plate counting. Si osserva soprattutto nella comunità KAN in presenza di antibiotico, in cui il ceppo LM supera gli altri due dopo le prime 70 ore. La spiegazione di questo fenomeno è dovuta alla presenza di cellule dei ceppi PM e PL che non esprimono la fluorescenza, che si stima siano il 2% delle cellule totali della comunità.

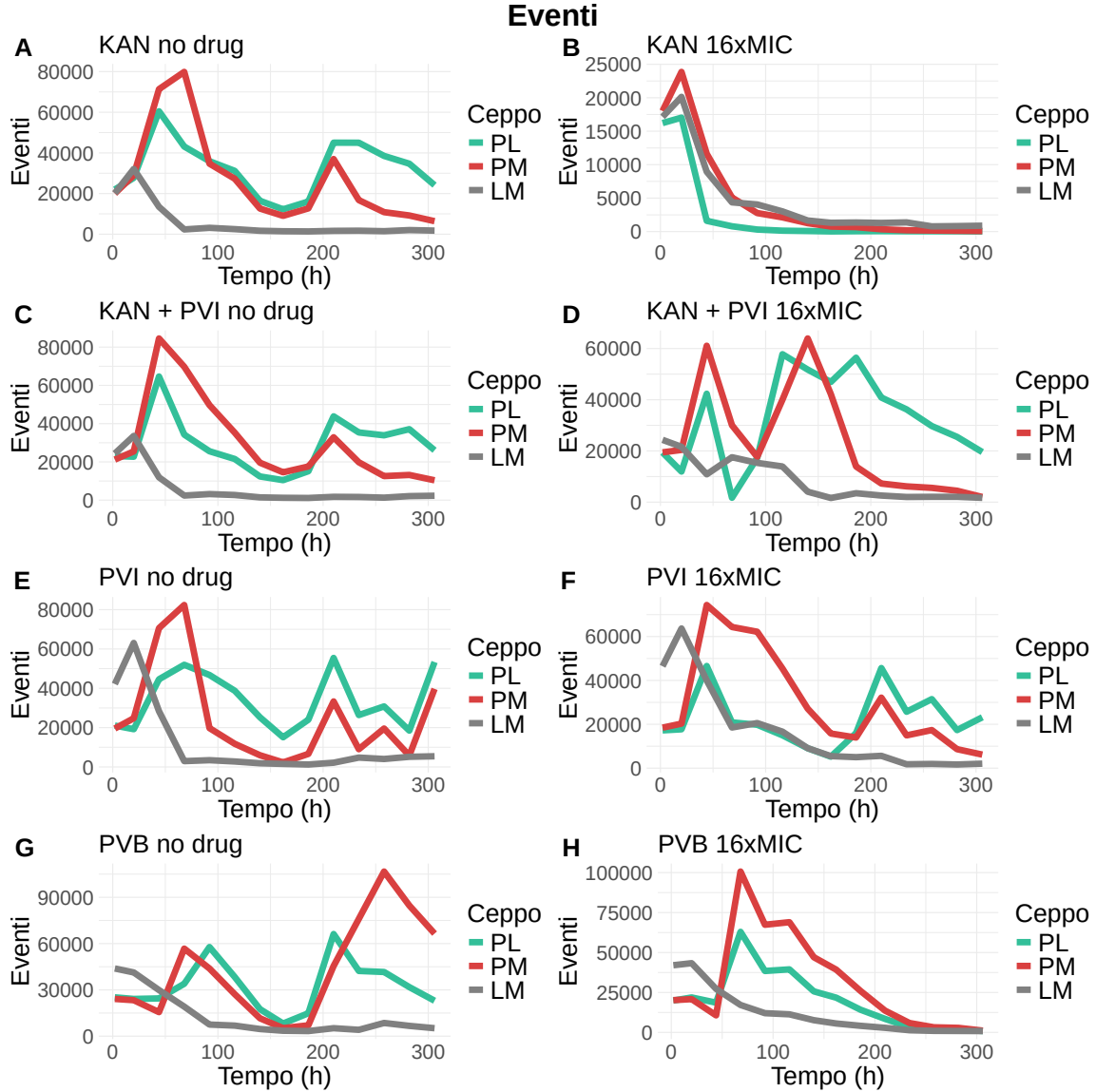


Figura 4.5: Eventi delle quattro comunità del primo esperimento in presenza e in assenza di ampicillina.

Nel secondo esperimento sono stati analizzati soltanto i dati di densità ottica e fluorescenza per i ceppi PM e PL dello strumento Chi.Bio. La comunità KAN senza antibiotico è stata studiata per 498 ore ed è stata confrontata con i dati ottenuti dal primo esperimento, osservando che continua ad oscillare anche a tempi successivi a 306 ore (Figura 4.6). Questo mostra la dinamica di comunità nel suo complesso, che in assenza di perturbazioni presenta un andamento oscillatorio.

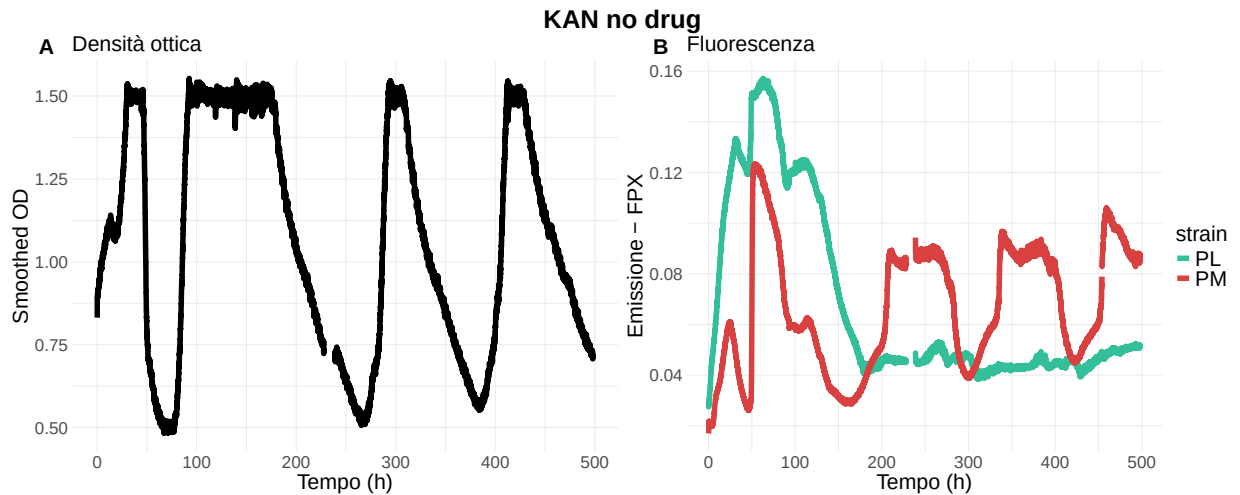


Figura 4.6: Densità ottica e fluorescenza della comunità KAN senza antibiotici del secondo esperimento.

Nel terzo esperimento sono stati analizzati le CFU dei due replicati (figure 4.7 e 4.8) di tutte le comunità in presenza e in assenza di colistina. I dati di densità ottica, fluorescenza ed eventi, presenti soltanto per il primo replicato, non sono mostrati poiché presentano un andamento analogo a quello delle CFU, che sono i dati che meglio esprimono la dinamica dei singoli ceppi delle comunità. La comunità KAN all'aumentare della dose di colistina dovrebbe mostrare un andamento che si avvicina sempre più alla morte cellulare. Nel primo replicato, per esempio, si vede che passa da avere due oscillazioni in assenza di antibiotico a un'unica oscillazione in presenza di una concentrazione 1xMIC di colistina. Quando però è in presenza di una concentrazione 2xMIC si nota in entrambi i replicati che l'andamento ricomincia ad oscillare; per spiegare questo fenomeno, è stato sequenziato l'intero genoma ed è stata trovata una mutazione genica nei ceppi PM e PL, che ne causa la sopravvivenza. Nel caso della comunità PN23, invece, i due replicati hanno andamenti diversi. Il primo replicato (Figura 4.7) mostra che, in assenza di antibiotico, la comunità oscilla molto meno rispetto a quella KAN, mentre aumentando le dosi di antibiotico la comunità passa da un'unica oscillazione (concentrazione 1xMIC di colistina) alla morte cellulare (concentrazione 2xMIC di colistina). Nel secondo replicato (Figura 4.8), la comunità in assenza di antibiotico ha lo stesso andamento osservato nel primo replicato, mentre aggiungendo la colistina si nota che, con una concentrazione 1xMIC di colistina, si ha il trasferimento del plasmide e la sopravvivenza della coltura, mentre con una concentrazione 2xMIC tutta la comunità muore. Infine, la comunità KAN + PN23 è stata testata solo in presenza di 2xMIC di colistina e nel primo replicato presenta delle oscillazioni, che sembrerebbero replicarsi anche nel secondo, che però è terminato prima di osservarle.

4 Risultati e discussione

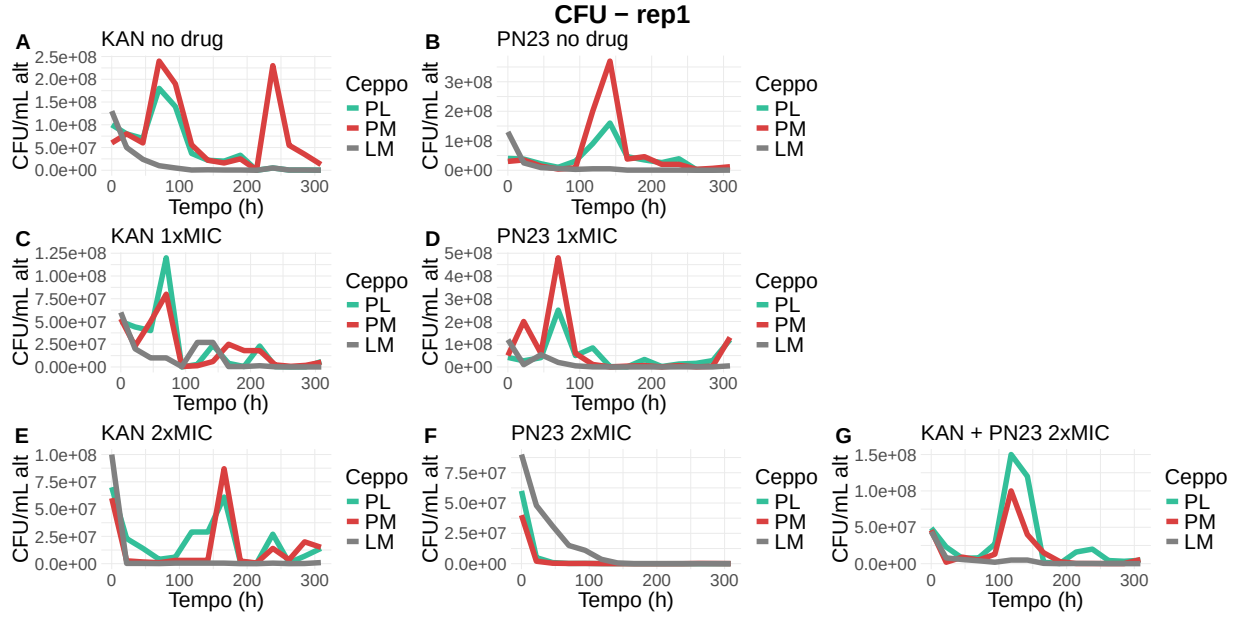


Figura 4.7: Plate counts (primo replicato) delle tre comunità del terzo esperimento in presenza e in assenza di colistina.

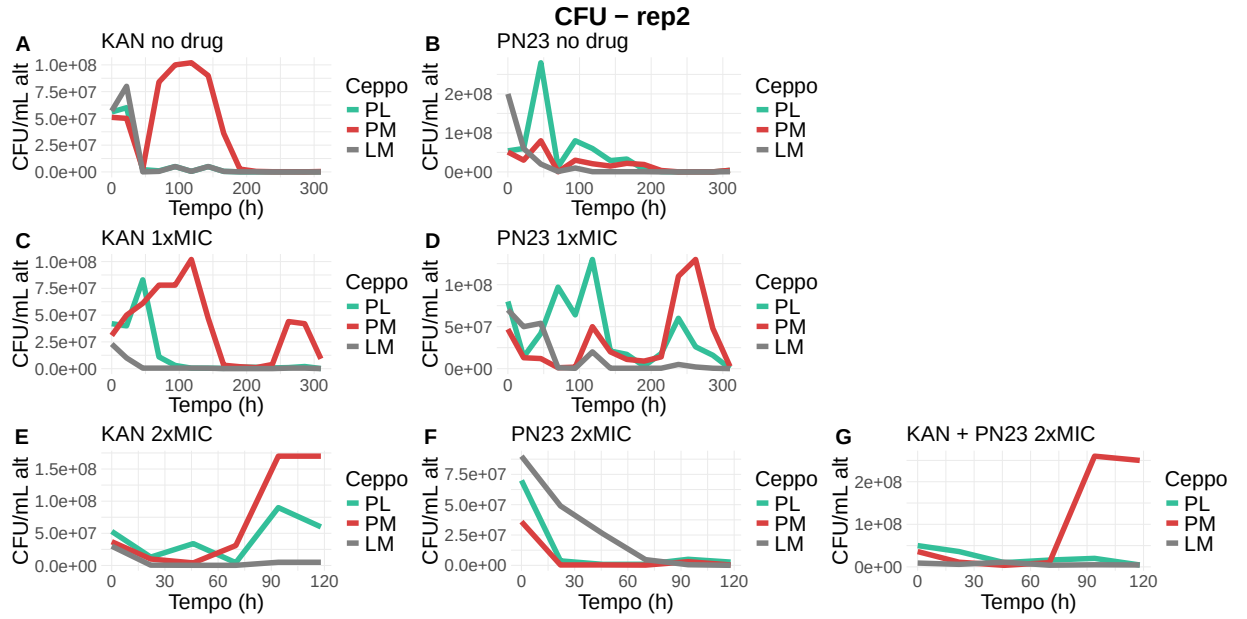


Figura 4.8: Plate counts (secondo replicato) delle tre comunità del terzo esperimento in presenza e in assenza di colistina.

Nell'ultimo esperimento vengono analizzati soltanto i dati di CFU delle comunità in assenza e in presenza di colistina, omettendo i dati di densità ottica, fluorescenza ed eventi per motivi analoghi a quelli spiegati per il terzo esperimento. Le comunità sono state coltivate per tempi superiori a quelli del terzo esperimento per osservare le dinamiche nel lungo periodo. La comunità KAN mostra un andamento oscillatorio che, all'aumentare della dose di colistina, diminuisce il numero di picchi fino a raggiungere la morte cellulare

quando l'antibiotico è in concentrazione 2xMIC, a differenza di quanto osservato nel terzo esperimento. La comunità PN23, invece, inizialmente oscilla, mentre, all'aggiungere di una concentrazione di 1xMIC, le oscillazioni diminuiscono, anche se la comunità sopravvive per il trasferimento del plasmide. Quando la colistina raggiunge una dose di 2xMIC, tutta la comunità muore.

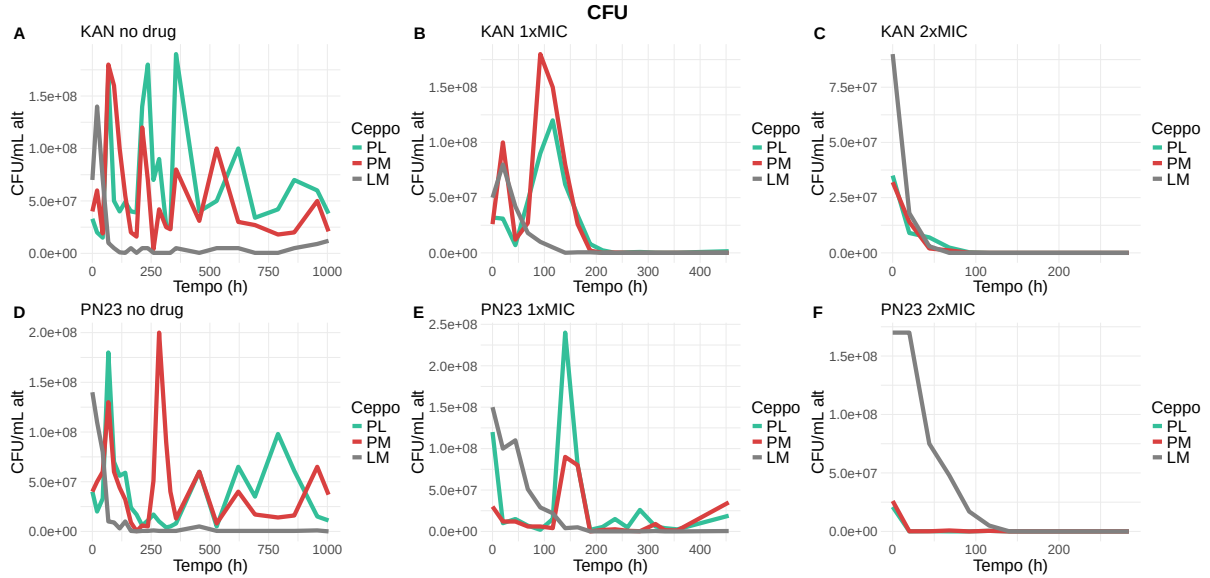


Figura 4.9: Plate counts delle tre comunità del quarto esperimento in presenza e in assenza di colistina.

4.3 Analisi della crescita dei ceppi sui terreni

In concomitanza con il primo esperimento, sono stati svolti alcuni esperimenti supplementari per studiare come cambia il tasso di crescita dei singoli ceppi all'interno di terreni arricchiti con aminoacidi. Sono stati analizzati separatamente in ogni terreno i ceppi LM delle comunità KAN, PVB e PVI, PM con e senza l'aggiunta del plasmide mScarlet, che esprime la fluorescenza rossa, e PL con e senza il plasmide gfp, che esprime la fluorescenza verde. I terreni utilizzati sono M9+PM, ossia M9 con l'aggiunta di fenilalanina e metionina, M9+PL, ossia M9 con l'aggiunta di fenilalanina e lisina, M9+LM, ossia M9 con l'aggiunta di lisina e metionina, e M9+Cas, ossia il terreno minimo con tutti gli aminoacidi aggiunti. I risultati di questa analisi sono mostrati nella heatmap 4.10, che è stata ottenuta su R (paragrafo 3.1) utilizzando i pacchetti *dplyr* per manipolare i dati e *pals* per ottenere la palette di colori espressa dalla funzione *parula*, che è la stessa utilizzata nelle heatmap ottenute con MATLAB (paragrafo 4.10). E' possibile notare che all'interno dei primi tre terreni i ceppi che crescono maggiormente sono quelli che riescono a produrre tutti gli aminoacidi perché ricevono dal terreno quelli mancanti. In particolare, nel primo terreno crescono principalmente PM e PM-mScarlet, nel secondo PL e PL-gfp e nel terzo LM-KAN, LM-PVB e LM-PVI. Nel terreno in cui sono presenti tutti gli aminoacidi, invece, tutti i ceppi crescono con tassi di crescita superiori a quelli osservati in M9+PM, M9+PL e M9+LM, poiché le specie batteriche, non dovendo sintetizzare autonomamente gli aminoacidi mancanti, risparmiano energia metabolica, che viene destinata alla crescita.

Dalla heatmap è possibile osservare anche l'impatto della presenza dei plasmidi sui tassi di crescita dei singoli ceppi. In particolare, tra le varianti del ceppo LM, LM-PVI è il ceppo che cresce maggiormente, seguito da LM-KAN e infine da LM-PVB. Questo dimostra che la presenza del plasmide PVI aumenta la fitness del batterio, mentre quella del plasmide PVB la diminuisce. Il ceppo PM, invece, ha un tasso di crescita maggiore in assenza del plasmide mScarlet, che impone un costo aggiuntivo alla fitness del batterio. Infine, il ceppo PL ha un tasso di crescita maggiore senza plasmide gfp soltanto nel terreno M9+PL, mentre nel terreno M9+Cas PL-gfp supera PL come tasso di crescita.

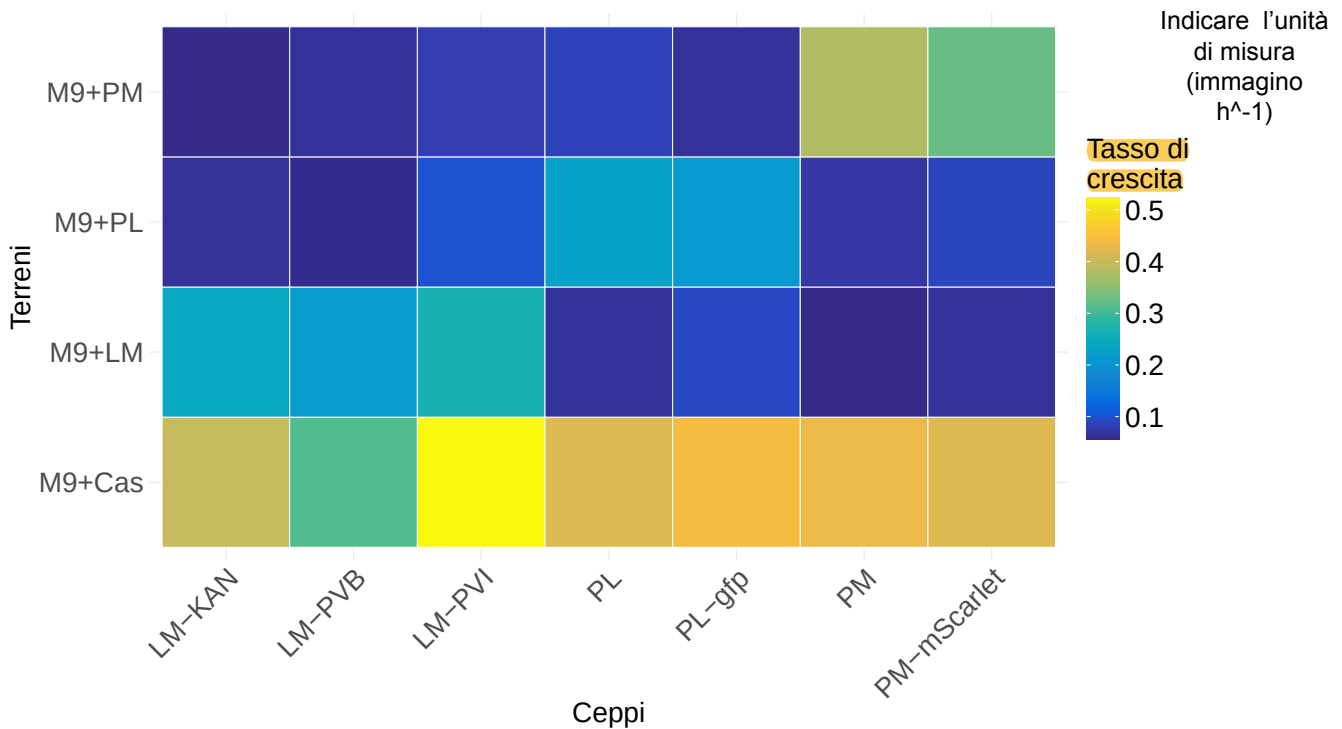


Figura 4.10: Heatmap dei tassi di crescita dei singoli ceppi all'interno di terreni minimi arricchiti da aminoacidi.

4.4 Analisi dati sul secondo esperimento

I dati misurati nel secondo esperimento sono densità ottica e fluorescenza per la comunità KAN senza antibiotico, mentre non si trovano misure di CFU, che sono il tipo di dato più adatto alla costruzione del modello, poiché forniscono informazioni sull'abbondanza e sulla dinamica di ogni ceppo nel tempo. In assenza dei dati di CFU del quarto esperimento, le misure di fluorescenza e densità ottica sono l'unico modo per predire l'andamento della comunità KAN senza antibiotico a tempi successivi a quelli del primo esperimento, per cui ci siamo chiesti se fosse possibile convertirle in CFU.

Abbiamo cercato una correlazione tra i dati di CFU e fluorescenze appartenenti al primo esperimento, in modo da poter utilizzare la stessa correlazione per calcolare le nuove CFU a partire dalle fluorescenze del secondo esperimento. Analizzando le correlazioni tra le CFU di PM e la fluorescenza rossa FP2 e tra le CFU di PL e la fluorescenza verde FP1 attraverso la funzione *curveFitter* di MATLAB (paragrafo 3.2.3), si ottiene che la migliore

è quella tra le CFU di PM e la fluorescenza FP2, visibile in figura 4.11a. La regressione è esponenziale e descritta dall'equazione 4.1, in cui i coefficienti a e b sono indicati nella tabella 4.1 con i relativi limiti superiore e inferiore dell'intervallo di confidenza al 95%. Le CFU del primo esperimento sono state normalizzate rispetto al loro ordine di grandezza (10^7) per facilitare i calcoli di correlazione.

$$f(x) = a \cdot e^{b \cdot x} \quad (4.1)$$

Coefficienti	Valore	IC 95% inferiore	IC 95% superiore
a	25.798	-2.0401	53.637
b	-14.946	-30.702	0.8105

Tabella 4.1: Coefficienti della regressione esponenziale e i loro intervalli di confidenza al 95%.

I parametri ottenuti che attestano la qualità del fit sono SSE, R-square, DFE, R-square aggiustato e RMSE, che valgono rispettivamente 312.25, 0.4026, 12, 0.3528 e 5.1011. Da questi parametri e dagli intervalli di confidenza di a e b è possibile affermare che la correlazione non è molto affidabile, nonostante sia la migliore tra quelle trovate. Inoltre, si nota che l'andamento della regressione è opposto rispetto a quello atteso: mentre ci si aspetterebbe che all'aumentare della fluorescenza aumentino anche le CFU, si osserva il contrario, con bassi valori di fluorescenza corrispondenti ad alti valori di CFU; i dati, inoltre, risultano abbastanza sparsi, a conferma della bassa affidabilità della correlazione. Sia la dispersione dei dati sia l'andamento invertito tra fluorescenza e CFU possono essere spiegati da motivi sperimentali, poiché i due tipi di dati provengono da due strumenti diversi, e dalla variabilità intrinseca presente nelle comunità microbiche.

Utilizzando l'equazione della regressione esponenziale 4.1 con i relativi coefficienti e considerando che x rappresenta la fluorescenza e y il numero di CFU, è possibile calcolare le CFU per il ceppo PM impiegando come fluorescenza i valori di FP2 misurati nel secondo esperimento. Il risultato è mostrato in figura 4.11b, da cui è possibile notare un andamento oscillatorio delle CFU del ceppo PM.

E' possibile calcolare anche le CFU del ceppo PL, conoscendo il rapporto delle fluorescenze e utilizzando le nuove CFU di PM nell'equazione 4.2. Il risultato è mostrato nella stessa figura 4.11b in cui è presente anche il ceppo PM, e mostra lo stesso andamento oscillatorio già riscontrato per PM.

$$CFU_{PL} = CFU_{PM} \cdot \left(\frac{FP1}{FP2} \right) \quad (4.2)$$

Poiché il ceppo LM non presenta fluorescenza, le sue CFU non possono essere ottenute direttamente dai dati sperimentali. Un metodo per calcolarle consiste nell'utilizzare la densità ottica. Ottenuta la correlazione tra numero di CFU totali e densità ottica, entrambi provenienti dal primo esperimento, è possibile calcolare le CFU totali del secondo esperimento a partire dai dati sperimentali di densità ottica. Si possono successivamente ottenere le CFU di LM sottraendo dal totale quelle di PM e PL appena calcolate.

E' presente una regressione esponenziale tra le CFU totali e la densità ottica, entrambi provenienti dal primo esperimento, come mostrato in figura 4.11c; la regressione segue

la stessa equazione 4.1 della regressione precedente, ma con i coefficienti indicati nella tabella 4.2. I parametri SSE, R-square, DFE, R-square aggiustato e RMSE ottenuti sono rispettivamente 1045.8, 0.3541, 12, 0.3002 e 9.3353. Anche in questo caso, la regressione non è molto affidabile, tuttavia l'andamento tra le CFU totali e la densità ottica è quello atteso, per cui all'aumentare delle CFU aumenta anche la densità ottica.

Coefficienti	Valore	IC 95% inferiore	IC 95% superiore
a	2.8639	-2.8018	8.5296
b	1.3823	-0.0449	2.8094

Tabella 4.2: Coefficienti della regressione esponenziale e i loro intervalli di confidenza al 95%.

Utilizzando l'equazione della regressione esponenziale 4.1 con i relativi coefficienti (dove x rappresenta la densità ottica e y il numero di CFU), è possibile ottenere il numero di CFU totali per il secondo esperimento a partire dai dati di densità ottica. Sottraendo dalle CFU totali quelle di PM e PL, si ottengono le CFU di LM, che sono mostrate in figura 4.11d. Dal confronto tra le CFU di LM e quelle di PM e PL, è possibile notare che LM ha un andamento opposto all'inizio, mentre successivamente rispecchia i picchi degli altri due ceppi. Il problema principale risiede nel fatto che alcuni valori delle CFU di LM risultano negativi, poiché in alcuni punti il numero di CFU totali è inferiore rispetto alla somma delle CFU calcolate di PM e PL. Per questo motivo, tale metodo di calcolo delle CFU di LM non è utilizzabile.

E' possibile considerare un altro approccio per ottenere le CFU di LM, partendo dalle densità ottiche dei singoli ceppi. Da esperimenti di calibrazione si sono ottenuti i dati di densità ottica di ogni ceppo e delle fluorescenze FP1 e FP2. Una volta calcolate le densità ottiche per i ceppi PM e PL a partire dalle loro fluorescenze, è possibile sottrarle dalla densità ottica misurata nel secondo esperimento e ottenere la densità ottica soltanto di LM, da cui si possono poi ricavare le CFU.

Poiché i dati di calibrazione sono troppo sparsi, non è possibile ottenere curve di correlazione tra densità ottica e fluorescenza; si può quindi utilizzare il rapporto tra le due grandezze per ottenere le nuove densità ottiche di PM e PL a partire dalle loro fluorescenze del secondo esperimento, con le equazioni 4.3 e 4.4.

$$OD\ PM = media\left(\frac{OD\ PM\ calibrazione}{FP2\ calibrazione}\right) \cdot FP2 \quad (4.3)$$

$$OD\ PL = media\left(\frac{OD\ PL\ calibrazione}{FP1\ calibrazione}\right) \cdot FP1 \quad (4.4)$$

La densità ottica di PM è visibile in figura 4.12a e quella di PL in figura 4.12b. In entrambi i casi, i valori di OD sono tra 0.2 e 1.2, in linea con l'intervallo atteso per i dati di densità ottica.

E' possibile ottenere la densità ottica di LM sottraendo dalla densità ottica del secondo esperimento quelle di PM e PL. Il risultato è visibile in figura 4.12c, in cui si nota che la densità ottica è negativa per gran parte del tempo. A partire da questa densità e considerando il rapporto tra CFU e OD ottenuto dai dati di calibrazione, si può utilizzare l'equazione 4.5 per ottenere le CFU di LM. Si nota che il coefficiente utilizzato per passare

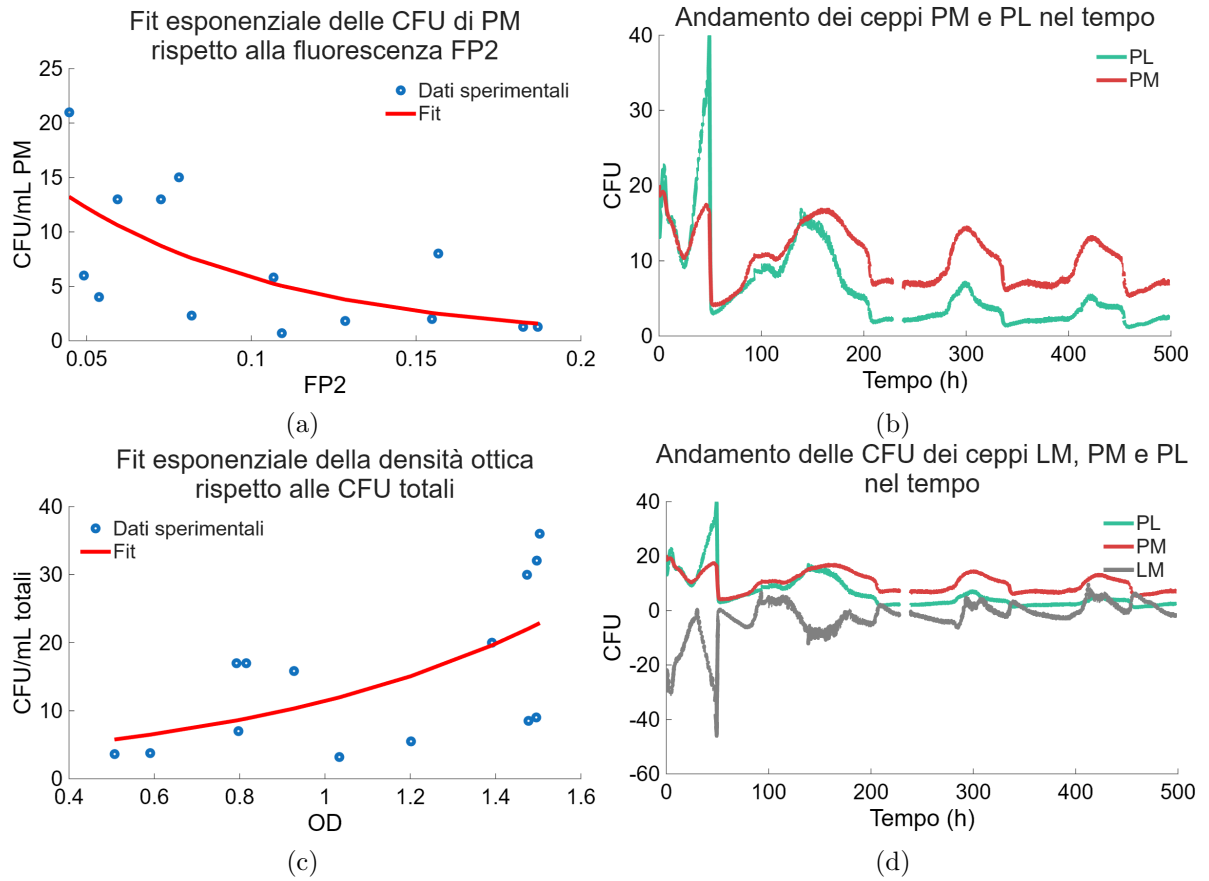


Figura 4.11: Regressioni esponenziali tra CFU di PM e fluorescenza FP2 [4.11a] e tra CFU totali e densità ottica [4.11c]. CFU di PM e PL [4.11b] e di LM, PM e PL [4.11d].

da OD a CFU non è la media dei rapporti tra i dati, come è avvenuto per PM e PL, ma è il rapporto tra le medie di CFU e OD. Questa scelta è dovuta al fatto che i dati provengono da due esperimenti di calibrazione diversi.

$$CFU \text{ LM} = \frac{\text{media}(CFU \text{ LM})}{\text{media}(OD \text{ LM})} \cdot OD \text{ LM} \quad (4.5)$$

Nella figura 4.12d, sono confrontate le CFU di LM ottenute con questo approccio e le CFU di PM e PL, calcolate rispettivamente tramite le equazioni 4.1 e 4.2. E' possibile osservare che, anche con questo secondo approccio, le CFU di LM presentano alcuni valori negativi, anche se rispecchiano maggiormente l'andamento oscillatorio di PM e PL. Questo potrebbe tuttavia essere un artefatto dei calcoli effettuati e non un dato reale.

In conclusione, non è possibile utilizzare i dati di fluorescenza e di densità ottica del secondo esperimento per ricavare le CFU dei singoli ceppi. I metodi proposti avrebbero potuto essere un modo per ottenere risultati, tuttavia si sono scontrati con diversi problemi. Il primo fra tutti riguarda l'inaffidabilità delle regressioni esponenziali dovuta ai dati sparsi e alla variabilità delle comunità batteriche, che non sempre seguono le previsioni attese. Un altro problema riguarda la comparabilità di dati provenienti da strumenti sperimentali diversi, che potrebbero non essere realmente confrontabili tra loro. L'ultimo problema è la presenza di CFU negative: poiché tramite il plate counting viene contato il numero di colonie (CFU), non è possibile avere un numero negativo; per questo motivo, l'anda-

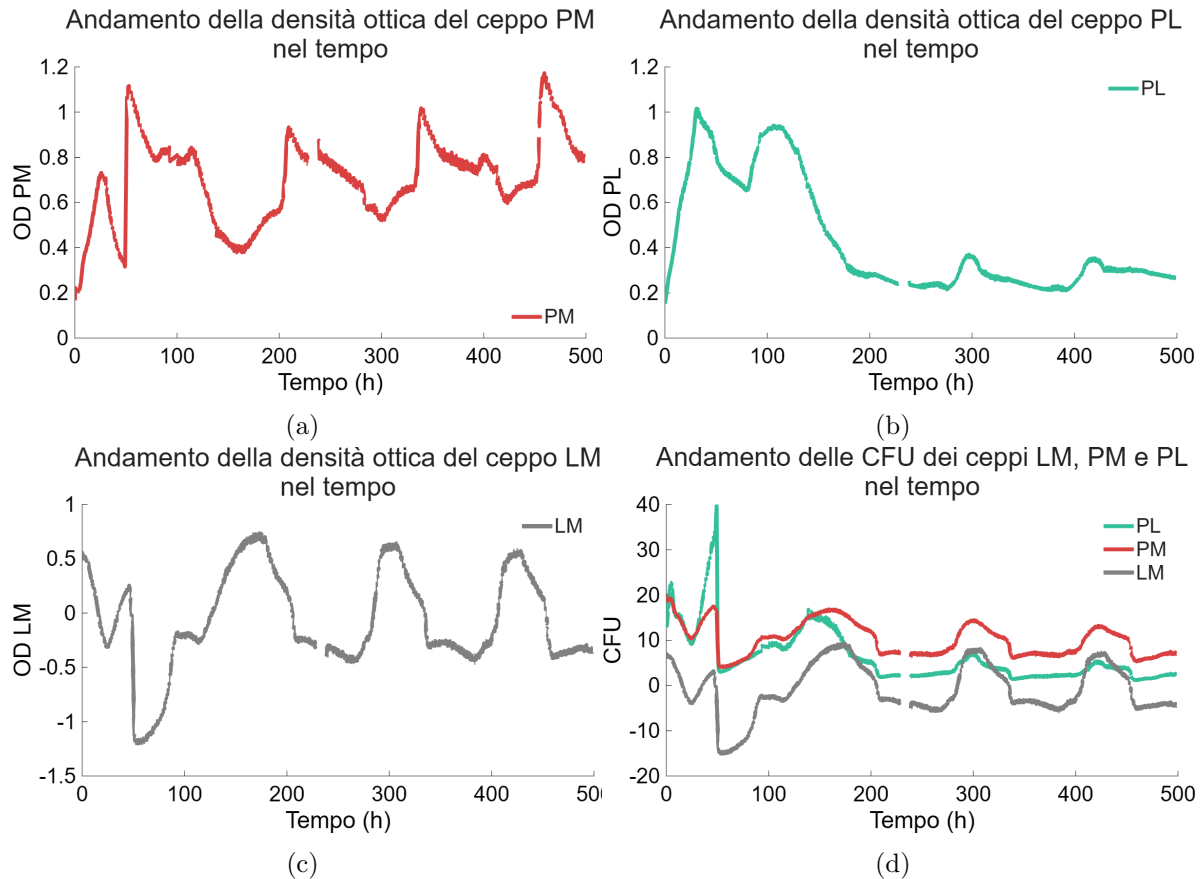


Figura 4.12: Densità ottiche per i singoli ceppi di PM [4.12a], PL [4.12b] e LM [4.12c], e di confronto delle CFU di LM, PM e PL [4.12d].

mento delle CFU con valori negativi non può essere utilizzato in un modello, in quanto non rispecchia la realtà. Il secondo esperimento viene utilizzato allora solo per osservare l'andamento della comunità a tempi sperimentali successivi a quelli del primo esperimento e per affermare che continua ad oscillare anche dopo 306 ore. E' stato successivamente superato dal quarto esperimento, che include anche la comunità KAN, monitorata per 1000 ore, e offre le misure di CFU dei singoli ceppi.

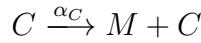
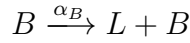
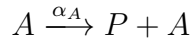
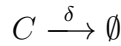
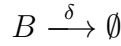
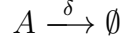
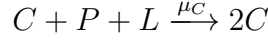
4.5 Modello matematico delle comunità in assenza di antibiotico

Il modello matematico che è stato utilizzato per descrivere la dinamica delle comunità è il modello Lotka-Volterra generalizzato. Inizialmente, si definiscono le equazioni chimiche 4.6 che esprimono le dinamiche di crescita e di morte dei singoli ceppi e di produzione degli aminoacidi. Successivamente, a partire dalle equazioni chimiche, si definiscono le equazioni differenziali 4.7, che sono utilizzate dal modello per simulare la dinamica di crescita batterica e la produzione degli aminoacidi da parte delle specie.

Di cosa? Degli aminoacidi?

Le equazioni chimiche 4.6 delle comunità che crescono in assenza di antibiotico sono nove. Le prime tre esprimono la **produzione** dei ceppi LM, PM e PL, che sono indicati rispettivamente come A, B e C. Ognuno dei ceppi, quando si combina con gli aminoacidi che non può produrre, cresce con un tasso di crescita μ diverso. La quarta, la quinta e la sesta equazione rappresentano la morte dei tre ceppi, che avviene a un tasso di morte

indicato da δ . E' importante precisare che il tasso di morte è lo stesso per tutti i ceppi e corrisponde al tasso di diluizione nella coltura. Infine, le ultime tre equazioni esprimono la produzione degli aminoacidi da parte dei tre ceppi con dei tassi di produzione α .



Le equazioni differenziali 4.7 ottenute a partire dalle equazioni chimiche sono sei. Le prime tre esprimono la dinamica di crescita dei ceppi A, B e C e sono composte da una parte che indica la loro produzione ($\mu_A[A][L][M]$ nel caso di A) e una parte che indica il loro consumo a causa della morte cellulare ($-\delta[A]$ nel caso di A). Le ultime tre rappresentano la variazione della concentrazione degli aminoacidi e sono formate da una parte che esprime il consumo dell'aminoacido, ossia il suo utilizzo da parte degli altri due ceppi, e da una parte che esprime la produzione da parte del ceppo specifico. Nel caso della lisina, viene consumata da A ($-\mu_A[A][L][M]$) e da C ($-\mu_C[C][P][L]$) e prodotta da B ($+\alpha_B[B]$).

$$\frac{d[A]}{dt} = \mu_A[A][L][M] - \delta[A] \quad (4.7)$$

$$\frac{d[B]}{dt} = \mu_B[B][P][M] - \delta[B]$$

$$\frac{d[C]}{dt} = \mu_C[C][P][L] - \delta[C]$$

$$\frac{d[L]}{dt} = -\mu_A[A][L][M] - \mu_C[C][P][L] + \alpha_B[B]$$

$$\frac{d[M]}{dt} = -\mu_A[A][L][M] - \mu_B[B][P][M] + \alpha_C[C]$$

$$\frac{d[P]}{dt} = -\mu_B[B][P][M] - \mu_C[C][P][L] + \alpha_A[A]$$

4.5.1 Fitting del modello utilizzando le equazioni differenziali

Le equazioni differenziali del modello presentano sette parametri: i tassi di crescita μ_A , μ_B e μ_C , i tassi di produzione degli aminoacidi α_A , α_B e α_C e il tasso di morte δ . Per ottenere

i valori di tali parametri, è necessario utilizzare la funzione di ottimizzazione *SIMPLEXL* (paragrafo 3.2.1), che, a partire da valori iniziali dei parametri scelti in modo casuale, permette di ottenere quelli migliori eseguendo un fitting sui dati sperimentali. Prima di svolgere il fitting, allora, è necessario decidere quali dati sperimentali utilizzare e come impostare i parametri iniziali e le opzioni della funzione *SIMPLEXL*.

I dati sperimentali utilizzati sono le CFU misurate attraverso il plate counting, poiché sono i dati che meglio rappresentano le fluttuazioni delle popolazioni dei ceppi. Per ottenere dei parametri robusti, che rappresentino realmente la dinamica della comunità, è necessario effettuare il fitting del modello su più replicati sperimentali. Poiché i due replicati della comunità KAN senza antibiotico del primo esperimento hanno durate temporali diverse, non è possibile utilizzarli insieme. Per risolvere il problema, si sono utilizzati come replicati quelli da 306 ore delle comunità KAN, KAN + PVI, PVI e PVB senza antibiotico, in modo da ottenere dei dati sperimentali che rappresentassero l'intera comunità. Il plasmide, che non viene trasferito in assenza di ampicillina, è presente soltanto nel ceppo LM, che dalle misure di CFU risulta sempre poco numeroso, mentre i ceppi PM e PL, che sono i più oscillanti, non lo contengono. Per questo motivo, si può considerare valida l'approssimazione di tutte le comunità come una unica senza antibiotico. Dai quattro replicati sono state ottenute le medie e le deviazioni standard e, dopo aver interpolato i dati, è stato eseguito il fitting del modello.

In alternativa, è possibile considerare come replicati della comunità KAN senza antibiotico queste tre serie di dati: il secondo replicato di KAN del primo esperimento, il primo replicato di KAN del terzo esperimento e il primo replicato di KAN del quarto esperimento. I motivi per cui non sono stati utilizzati questi tre replicati fin dall'inizio sono tre. Il primo è legato a motivi temporali, poiché il terzo e il quarto esperimento sono stati svolti alcuni mesi dopo rispetto al primo. Il secondo riguarda i tempi dei dati sperimentali, che variano di qualche ora tra i vari replicati. Questo secondo problema può essere in parte superato considerando che le misure sono state effettuate una volta al giorno, per cui uno spostamento di due o tre ore rispetto all'orario segnato nei dati non modifica la misura del numero di CFU. Il terzo motivo è che il primo replicato del quarto esperimento ha una durata molto superiore rispetto agli altri due, per cui sono stati considerati solo i dati fino a 306 ore. Anche per questi replicati, sono state ottenute le medie e le deviazioni standard e, dopo aver interpolato i dati, è stato eseguito il fitting del modello.

Prima di eseguire il fitting, sono state impostate le opzioni della funzione *foptions* (paragrafo 3.2.1) da introdurre nella funzione *SIMPLEXL*, che sono le *options(2)*, *options(3)* e *options(14)*, dai valori rispettivamente di 10^{-8} , 10^{-8} e 10^3 . Inoltre, i parametri iniziali inseriti sono stati $p0 = [0.0004, 0.0005, 0.0005, 0.7759, 0.0520, 0.1324, 0.0702]$. Dati questi parametri e opzioni, utilizzando la funzione *SIMPLEXL* sulle medie dei replicati ottenuti dal primo esperimento, si ottengono, come parametri migliori, i valori 0.0003, 0.0005, 0.0004, 0.7645, 0.0567, 0.1202 e 0.0692, che in ordine rappresentano μ_A , μ_B , μ_C , α_A , α_B , α_C e δ . Il fitting del modello sui dati, utilizzando questi parametri calcolati, è mostrato in figura 4.13a.

E' stato eseguito il fitting anche sulle medie dei tre replicati della comunità KAN senza antibiotico, utilizzando le stesse opzioni di *foptions*, ma modificando i parametri iniziali in $p0 = [0.0023, 0.0028, 0.0022, 0.8928, 0.0438, 0.0942, 0.0591]$. I parametri ottenuti tramite *SIMPLEXL* sono 0.0006, 0.0018, 0.0027, 0.3958, 0.0362, 0.0644 e 0.0398, ed è possibile osservare il fitting del modello in figura 4.13b.

Dalle analisi sperimentali, sono stati ottenuti i seguenti tassi di crescita in terreni M9 con lisina, metionina e fenilalanina: 0.029 per LM, 0.28 per PM e 0.117 per PL. Volendo integrare questi dati all'interno del modello, si sono fissati i parametri μ_A , μ_B e μ_C , assegnando loro, in ordine, i valori 0.00029, 0.0028 e 0.00117. Eseguendo l'ottimizzazione di *SIMPLEXL* con le stesse opzioni precedenti e i valori iniziali $p_0 = [0.8511, 0.0552, 0.1123, 0.0647]$, si sono ottenuti come parametri migliori 0.4054, 0.0365, 0.0693 e 0.0389, che corrispondono rispettivamente a α_A , α_B , α_C e δ . Il fitting derivato da questi parametri è in figura 4.13c.

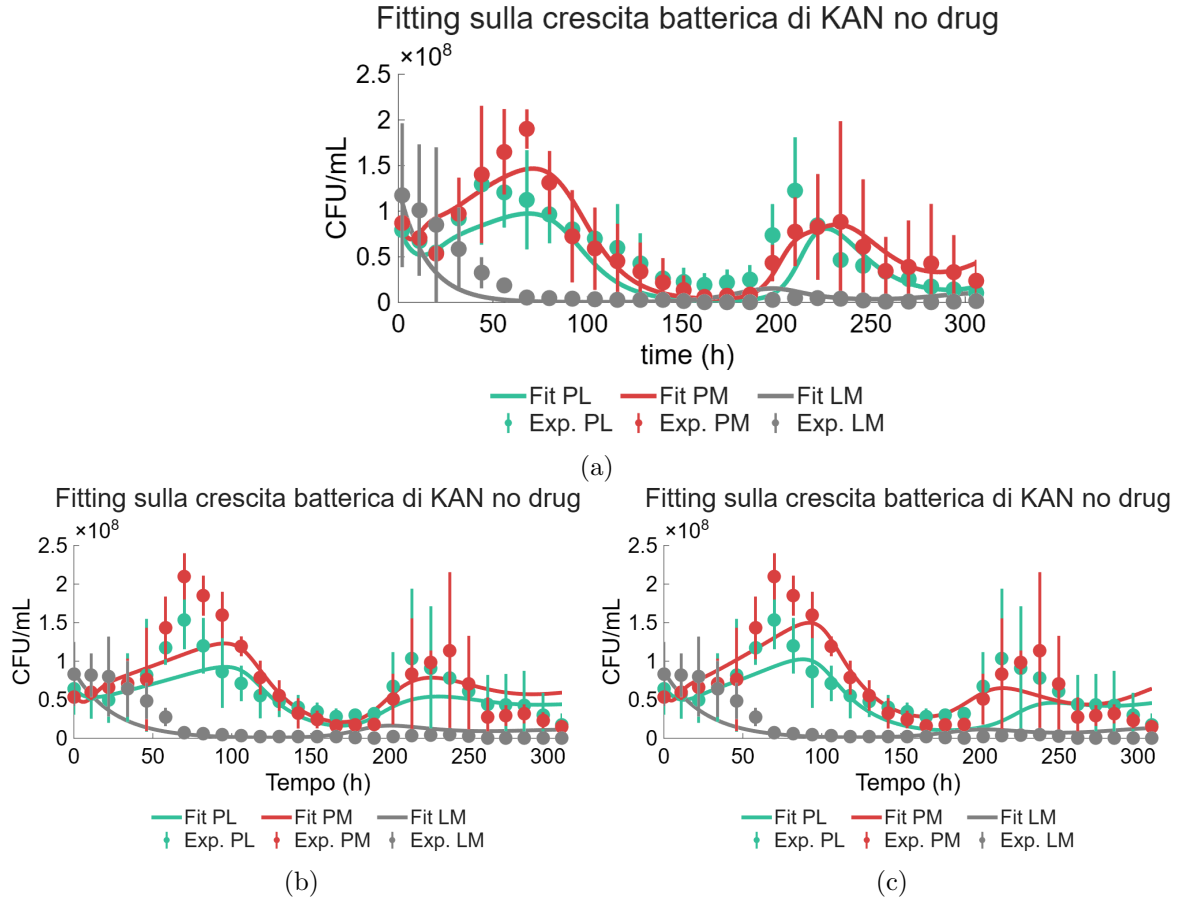


Figura 4.13: Fitting a partire dai dati delle comunità KAN, KAN+PVI, PVI e PVB senza ampicillina del primo esperimento [4.13a] e dai dati delle comunità KAN ottimizzando tutti i parametri [4.13b] e ottimizzando soltanto i tassi di produzione degli aminoacidi e di morte [4.13c].

E' possibile confrontare i valori dei parametri ottenuti nei tre fitting nella tabella 4.3. Dai tassi di crescita ottenuti sperimentalmente è stato notato che il ceppo PM cresce 9.655 volte di più rispetto a LM, mentre il ceppo PL 4.034 volte di più rispetto a LM. Questi rapporti non sono rispettati negli altri due fitting, come è possibile vedere nella tabella 4.4. Questo può essere spiegato dal fatto che i tassi di crescita sperimentali sono misurati in monoculture dei singoli ceppi, per cui in presenza di più ceppi non rimangono uguali, ma tendono a diminuire. Dallo stesso esperimento, è stato osservato che il tasso di crescita di tutti i ceppi quando sono messi insieme in un terreno minimo con lisina, metionina e fenilalanina misura 0.119, ossia un valore a metà tra quelli dei singoli ceppi. Per questo motivo, la crescita del ceppo PM viene limitata dalla presenza degli altri due ceppi e il suo

Para- metri	fit comunità plasmide	fit comunità KAN e tutti i parametri	fit comunità KAN e quattro parametri
μ_A	0.0003	0.0006	0.00029
μ_B	0.0005	0.0018	0.0028
μ_C	0.0004	0.0027	0.00117
α_A	0.7645	0.3958	0.4054
α_B	0.0567	0.0362	0.0365
α_C	0.1202	0.0644	0.0693
δ	0.0692	0.0398	0.0389

Tabella 4.3: Parametri ottenuti da *SIMPLEXL* per le comunità KAN, KAN+PVI, PVI e PVB senza ampicillina del primo esperimento, indicate come *fit comunità plasmide* e le comunità KAN ottimizzando tutti i parametri, indicate come *fit comunità KAN e tutti i parametri* e ottimizzando soltanto i tassi di produzione degli aminoacidi e di morte, indicate come *fit comunità KAN e quattro parametri*.

tasso di crescita si avvicina a quello degli altri. E' possibile osservare anche che il tasso di crescita di LM trovato nel primo fitting è quasi uguale a quello sperimentale una volta riportato allo stesso ordine di grandezza, essendo diminuito di due ordini di grandezza, anche se non rispecchia i rapporti con i tassi di crescita degli altri due ceppi.

Confrontando gli altri parametri calcolati (tabella 4.3), si nota che le due serie ottenute dal fitting della comunità KAN hanno valori molto simili, che tuttavia non sono troppo distanti dalla prima serie di parametri. I tassi di produzione degli aminoacidi di LM sono in tutti i casi maggiori rispetto a quelli di PM e PL e il primo fitting ha un valore quasi doppio rispetto agli altri due, così come avviene anche per il tasso di morte. I tassi di produzione di PM si avvicinano tra loro e quello del primo fitting non è molto più grande. Il tasso di produzione di PL, invece, è il doppio nel primo fitting rispetto agli altri due. Quello che si può notare è che nel primo fitting i valori di tutti i parametri, ad esclusione dei tassi di crescita, sono maggiori rispetto a quelli degli altri due, per cui vengono prodotti più aminoacidi, ma le cellule muoiono anche in quantità superiore, andando a bilanciare l'effetto.

Dal confronto tra i tre fitting, si considera come migliore il primo, poiché è quello che segue meglio l'andamento dei suoi dati sperimentali, che non utilizza dati con tempi modificati e che ha i tassi di crescita più simili tra loro. Quest'ultimo motivo, in particolare, è in linea con l'osservazione sperimentale che il tasso di crescita di tutti i ceppi tende a raggiungere un valore intermedio rispetto ai valori ottenuti per i singoli ceppi. Anche dalle simulazioni sul lungo periodo, presenti nel paragrafo 4.5.3, risulta essere il fitting che meglio rappresenta la dinamica di popolazione delle comunità.

4.5.2 Misura dell'intervallo di confidenza al 95%

Una volta ottenuti i parametri migliori, è necessario calcolare la loro affidabilità attraverso la misura dell'intervallo di confidenza al 95%. Tale intervallo è ricavato dalla funzione

Rapporti	fit comunità plasmide	fit comunità KAN e tutti i parametri	fit comunità KAN e quattro parametri
μ_A/μ_A	1	1	1
μ_B/μ_A	1.7	3	9.655
μ_C/μ_A	1.3	4.5	4.034

Tabella 4.4: Proporzioni tra i tassi di crescita per le comunità KAN, KAN+PVI, PVI e PVB senza ampicillina del primo esperimento, indicate come *fit comunità plasmide* e le comunità KAN ottimizzando tutti i parametri, indicate come *fit comunità KAN e tutti i parametri* e ottimizzando soltanto i tassi di produzione degli aminoacidi e di morte, indicate come *fit comunità KAN e quattro parametri*.

nlparci di MATLAB (paragrafo 3.2.2), che richiede come argomenti i parametri calcolati, il vettore dei residui, che è calcolato direttamente dalla *objective function* (paragrafo 3.2.1) e rappresenta la differenza tra il valore predetto dal modello e quello sperimentale, e lo Jacobiano. Esistono funzioni di MATLAB, come *nlinfit*, che producono direttamente, come output, i valori richiesti in input da *nlparci*, tuttavia non utilizzano la funzione *SIMPLEXL* (paragrafo 3.2.1) per ottenere i parametri ottimizzati. Per questo motivo, non possono essere utilizzate e lo Jacobiano deve essere ricavato in altri modi. Sapendo che è una matrice che esprime quanto si modifica il residuo al variare dei parametri, è possibile perturbare un parametro alla volta, modificandolo leggermente in positivo e in negativo, ed eseguire nuovamente la *objective function* per ottenere un nuovo vettore di residui per ogni perturbazione. Dai nuovi residui, viene calcolata la derivata parziale e inserita all'interno dello Jacobiano (script di MATLAB presente nei Materiali Supplementari [6.6]). Una volta ottenuti tutti i valori da inserire nella funzione *nlparci*, è possibile ricavare le misure degli intervalli di confidenza al 95% per i parametri calcolati dal fitting delle comunità senza antibiotico del primo esperimento, mostrate nella tabella 4.5. E' possibile ricavare nello stesso modo anche gli intervalli di confidenza dei parametri calcolati dai fitting delle comunità KAN, tuttavia l'attenzione si è concentrata soltanto sui parametri ottenuti dal primo fitting, che è ritenuto il migliore. Osservando gli intervalli di confidenza, che non sono molto larghi, è possibile dedurre che i parametri sono affidabili.

4.5.3 Simulazione del modello nel lungo periodo e confronto con i dati sperimentali

Una volta determinato che i parametri ottimizzati sono affidabili, è possibile inserirli all'interno del modello e simulare la dinamica di crescita della comunità nel lungo periodo, per osservare l'andamento dei singoli ceppi. Una comunità batterica nel lungo periodo può comportarsi in tre modi: raggiungere lo steady state, ossia uno stato in cui le abbondanze dei singoli ceppi non si modificano più, ma rimangono costanti; oscillare con oscillazioni smorzate, che diminuiscono di ampiezza fino a smettere di oscillare; oscillare con oscillazioni persistenti, che mantengono la stessa ampiezza e non smettono di oscillare nel tempo. Per comprendere il comportamento della comunità nel lungo periodo, si sono osservati i dati della comunità KAN senza antibiotico del quarto esperimento, che si estendono per 1000 ore, ed è stato possibile affermare che la comunità KAN oscilla con oscillazioni smorzate: mentre i primi tre picchi hanno un'ampiezza simile, il quarto

Parametri	Valore	IC 95% inferiore	IC 95% superiore
μ_A	0.000317126	0.000219364	0.000414888
μ_B	0.000507074	0.000253461	0.000760687
μ_C	0.000434281	0.000260793	0.000607768
α_A	0.764541	0.593188	0.935894
α_B	0.0567025	0.0470679	0.0663371
α_C	0.120152	0.0996954	0.140608
δ	0.0692179	0.0636167	0.0748191

Tabella 4.5: Valori migliori ottenuti per i parametri e gli intervalli di confidenza al 95% inferiore e superiore.

e il quinto diminuiscono di ampiezza. Allungando la simulazione del modello a 1000 ore e utilizzando i parametri ottenuti dal primo fitting, è possibile notare che le oscillazioni tendono a smorzarsi e ad apparire qualitativamente simili a quelle sperimentali (Figura 4.14). Poiché l'andamento del modello e quello dei dati sperimentali non sono gli stessi, non è possibile ottenere previsioni quantitative dalle simulazioni, tuttavia si può predire in modo qualitativo il comportamento della comunità e il suo andamento ad oscillazioni smorzate.

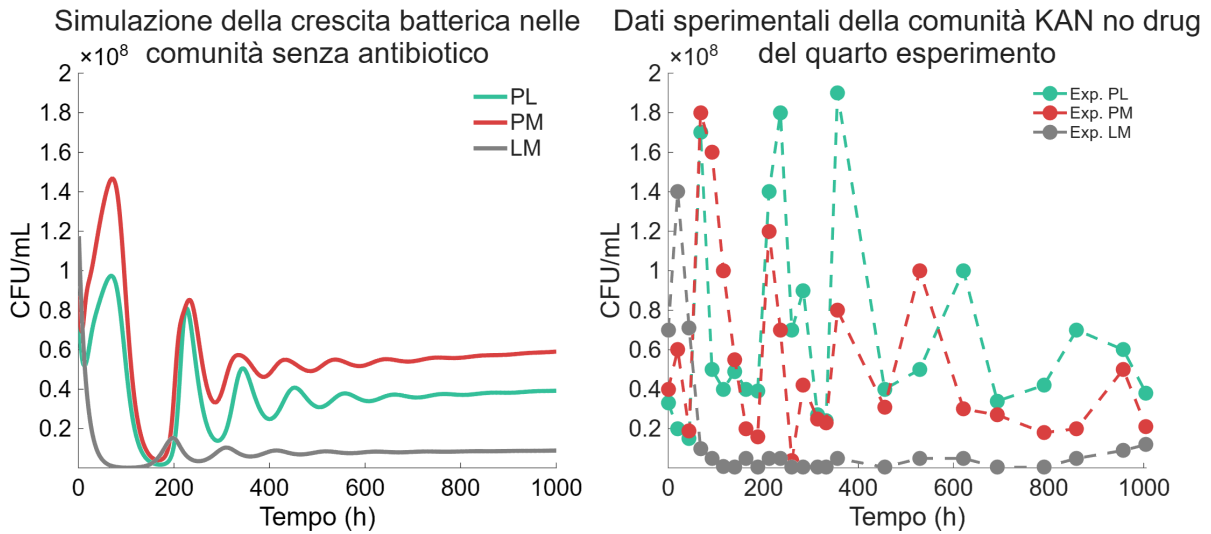


Figura 4.14: Comparazione della simulazione del modello (a sinistra), che utilizza i parametri del primo fitting, e dei dati sperimentali a 1000 ore del quarto esperimento (a destra).

E' possibile osservare anche le simulazioni nel lungo periodo dei modelli con i parametri ottenuti dal secondo e dal terzo fitting, ossia quello della comunità KAN ottimizzata per tutti i parametri (Figura 4.15a) e quello della comunità ottimizzata solo per i tassi di produzione degli aminoacidi e di morte (Figura 4.15b). La simulazione del secondo fitting, dopo le oscillazioni iniziali, si appiattisce e cresce in modo lento, senza oscillazioni e senza raggiungere uno steady state. La simulazione del terzo fitting, invece, cresce in

modo esponenziale e non oscilla ulteriormente. Dalle due simulazioni è possibile confermare che i parametri ottenuti dal primo fitting sono quelli migliori, poiché, anche se non permettono di ottenere predizioni quantitative, producono simulazioni qualitativamente simili all'andamento dei dati reali e mostrano il comportamento a lungo termine delle comunità batteriche in assenza di antibiotico.

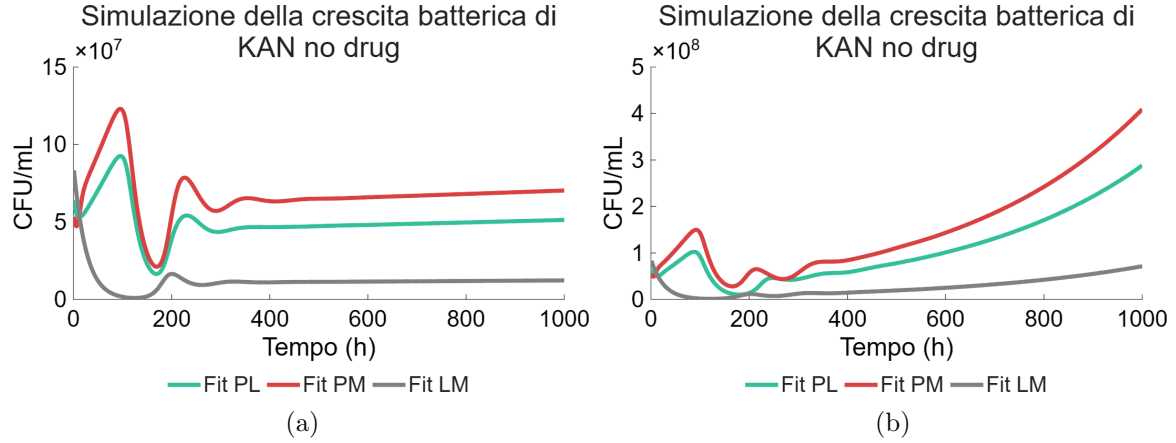
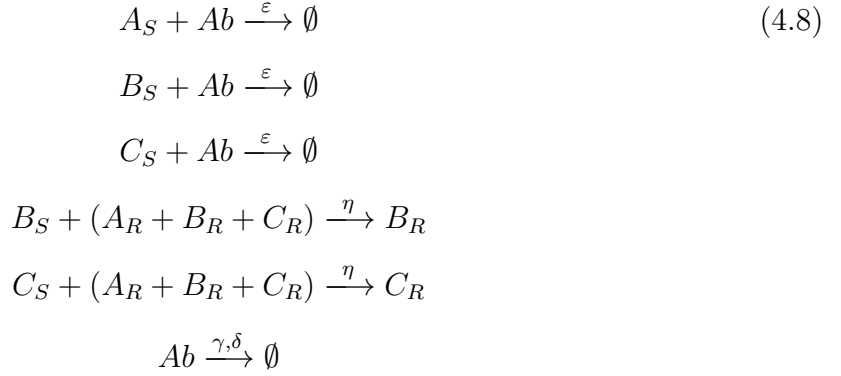


Figura 4.15: Simulazione dei modelli che utilizzano i parametri ottimizzati sui dati della comunità KAN. Il grafico 4.14 usa i parametri del secondo fitting, mentre quello 4.15b i parametri del terzo fitting.

4.6 Modello matematico delle comunità in presenza di antibiotico

Per descrivere la dinamica delle comunità in presenza dell'antibiotico, è necessario introdurre dei parametri in più all'interno delle equazioni del modello definito per le comunità senza antibiotico, in modo da esprimere l'effetto della presenza dell'antibiotico sui ceppi sensibili e della coniugazione tra ceppi resistenti e sensibili. Le equazioni chimiche 4.8 aggiunte al modello sono sei. Le prime tre esprimono la morte dei ceppi sensibili A, B e C a causa dell'antibiotico Ab e presentano un tasso di morte indotta dall'antibiotico ε . La quarta e la quinta equazione rappresentano la trasformazione dei ceppi B e C da sensibili a resistenti attraverso la coniugazione con i ceppi resistenti e hanno un tasso di coniugazione η . L'ultima equazione esprime la degradazione dell'antibiotico per motivi chimici (tasso di degradazione γ) e per diluizione (tasso di diluizione δ). Il tasso di diluizione può anche essere considerato diverso rispetto al tasso di morte dei tre ceppi, ma in questo caso l'abbiamo considerato uguale perché assumiamo che la morte cellulare avvenga solamente per diluizione.



Le equazioni differenziali 4.9, oltre ad integrare i termini espressi dalle equazioni chimiche, vengono modificate raddoppiando quelle riferite alla crescita dei ceppi, in modo da esprimere le dinamiche dei ceppi sensibili e resistenti separatamente. I ceppi indicati con S sono sensibili, mentre quelli con R resistenti. I termini come $-\varepsilon \cdot Ab \cdot A_S$ si riferiscono all'effetto dell'antibiotico sui ceppi sensibili, che è dovuto sia alla concentrazione dell'antibiotico che alla densità cellulare, mentre i termini come $-\eta \cdot B_S \cdot (A_R + B_R + C_R)$ si riferiscono alla trasformazione dei batteri da sensibili a resistenti, infatti questi termini diminuiscono la concentrazione di cellule sensibili e aumentano quella delle cellule resistenti. I ceppi resistenti hanno un tasso di crescita μ^R , che è diverso rispetto a quello dei ceppi sensibili a causa della presenza del plasmide, che può aumentare o diminuire la crescita batterica. Nel caso del plasmide PVI, come visto in precedenza, il tasso di crescita aumenta, per cui si ha un miglioramento della fitness portato dalla presenza del plasmide. Gli aminoacidi vengono prodotti sia dai ceppi sensibili che dai resistenti, ma poiché si trova poca differenza tra i tassi di crescita dei due tipi, si utilizza solamente quello dei sensibili come approssimazione.

$$\begin{aligned}
 \frac{dA_S}{dt} &= \mu_A \cdot A \cdot L \cdot M - \delta \cdot A_S - \varepsilon \cdot Ab \cdot A_S \\
 \frac{dA_R}{dt} &= \mu_A^R \cdot A_R \cdot L \cdot M - \delta \cdot A_R \\
 \frac{dB_S}{dt} &= \mu_B \cdot B_S \cdot P \cdot M - \delta \cdot B_S - \varepsilon \cdot Ab \cdot B_S - \eta \cdot B_S \cdot (A_R + B_R + C_R) \\
 \frac{dB_R}{dt} &= \mu_B^R \cdot B_R \cdot P \cdot M - \delta \cdot B_R + \eta \cdot B_S \cdot (A_R + B_R + C_R) \\
 \frac{dC_S}{dt} &= \mu_C \cdot C_S \cdot P \cdot L - \delta \cdot C_S - \varepsilon \cdot Ab \cdot C_S - \eta \cdot C_S \cdot (A_R + B_R + C_R) \\
 \frac{dC_R}{dt} &= \mu_C^R \cdot C_R \cdot P \cdot L - \delta \cdot C_R + \eta \cdot C_S \cdot (A_R + B_R + C_R) \\
 \frac{dP}{dt} &= -\mu_B \cdot (B_S + B_R) \cdot P \cdot M - \mu_C \cdot (C_S + C_R) \cdot P \cdot L + \alpha_A \cdot (A_S + A_R) \\
 \frac{dL}{dt} &= -\mu_A \cdot (A_S + A_R) \cdot L \cdot M - \mu_C \cdot (C_S + C_R) \cdot P \cdot L + \alpha_B \cdot (B_S + B_R) \\
 \frac{dM}{dt} &= -\mu_A \cdot (A_S + A_R) \cdot L \cdot M - \mu_B \cdot (B_S + B_R) \cdot P \cdot M + \alpha_C \cdot (C_S + C_R) \\
 \frac{dAb}{dt} &= -(\gamma + \delta) \cdot Ab
 \end{aligned} \tag{4.9}$$

L'ultima equazione esprime la dinamica dell'antibiotico, che normalmente segue un andamento di decadimento esponenziale espresso dall'equazione 4.10, in cui Ab_0 è la concentrazione iniziale di antibiotico e γ è il tasso di degradazione.

$$Ab(t) = Ab_0 \cdot e^{-\gamma t} \quad (4.10)$$

Le equazioni a partire da 4.9 permettono di modellare tutti i tipi di comunità con ampicillina e colistina, aggiungendo o azzerando i termini riferiti all'effetto dell'antibiotico e alla coniugazione. La comunità KAN in presenza di ampicillina o di colistina è modellata soltanto dalle equazioni che si riferiscono ai termini sensibili, non essendo presenti i resistenti. Per questo motivo, A_R , B_R , C_R e γ sono nulli. Le comunità KAN + PVI, PVI, KAN + PN23 e PN23 sono modellate con tutte le equazioni, tranne quella riferita al ceppo A sensibile, che non è presente in nessuna, infatti A_S è nullo. Nei casi delle comunità KAN + PVI e KAN + PN23 teoricamente i ceppi LM resistenti sono aggiunti dopo due ore rispetto all'inizio, per cui dovrebbero essere presenti anche i ceppi LM sensibili e dovrebbe essere aggiunta un'ulteriore equazione differenziale che rappresenti i ceppi sensibili che diventano resistenti. Poiché però la parte di ceppi sensibili non è particolarmente rilevante e il modello vuole simulare in modo solamente qualitativo l'andamento delle comunità, per semplicità le comunità KAN + PVI e KAN + PN23 vengono approssimate alle rispettive comunità PVI e PN23, che presentano i ceppi resistenti fin dall'inizio. La comunità PVB, invece, è modellata con le equazioni che si riferiscono ai ceppi A resistente e B e C sensibili. Non essendoci la possibilità di trasferire il plasmide, γ è nullo e lo stesso vale per A_S , B_R e C_R , che non possono diventare resistenti.

Nel caso in cui ampicillina e colistina venissero aggiunti costantemente al terreno di coltura nelle concentrazioni rispettivamente di 16xMIC e 1xMIC o 2xMIC, dovrebbe essere aggiunta l'equazione chimica 4.11, in cui la dose di antibiotico iniziale viene aggiunta dall'esterno con il tasso di diluizione δ .

$$\emptyset \xrightarrow{\delta \cdot Ab_{init}} Ab \quad (4.11)$$

Inoltre, l'ultima equazione differenziale del modello 4.9 dovrebbe essere sostituita dall'equazione 4.12, che integra il termine di ingresso dell'antibiotico.

$$\frac{dAb}{dt} = \delta \cdot Ab_{init} - (\gamma + \delta) \cdot Ab \quad (4.12)$$

4.7 Studio della degradazione di ampicillina e colistina

L'ampicillina è un antibiotico β -lattamico che appartiene alla classe delle penicilline, di cui è il componente più stabile agli acidi (paragrafo 3.4.1). E' un composto anfotero che possiede un anello β -lattamico, un anello tiazolidinico, un gruppo carbossilico e una catena laterale acilica (Figura 4.16). A pH 7 viene degradato principalmente per l'idrolisi dell'anello β -lattamico, che produce un sottoprodotto con attività antimicrobica ridotta. Il tasso di degradazione chimica γ del nostro modello, che in questo caso corrisponde al tasso di idrolisi, è influenzato dalle condizioni ambientali, come temperatura e pH. Alla temperatura di 37 °C e a pH 7, è possibile ottenere γ a partire dall'equazione 4.13, in cui k_h è il tasso di degradazione espresso in d^{-1} , E_a è l'energia di attivazione sperimentale,

che a pH 7 vale 81.2 kJ/mol, R è la costante universale dei gas ($8.3141 \text{ JK}^{-1}\text{mol}^{-1}$), T è la temperatura assoluta a 37°C (310.15 K) e A è il coefficiente di Arrhenius, il cui logaritmo naturale a pH 7 vale 28.92 d^{-1} [25].

$$\ln(k_h) = -\frac{E_a}{RT} + \ln(A) \quad (4.13)$$

Dall'equazione 4.13 si ottiene che k_h vale 0.0765 d^{-1} , che può essere trasformato nel coefficiente di degradazione γ del modello quando viene convertito in ore, ottenendo un valore di 0.0032 h^{-1} . Hou et al. [21] hanno osservato che il coefficiente di idrolisi può essere accelerato quando sono presenti dei fosfati nel terreno, come nel caso di M9. Poiché però il modello è soltanto qualitativo e per calcolare in modo accurato γ sono necessarie diverse analisi sperimentali, il tasso di accelerazione dovuto ai fosfati è stato trascurato, considerando come γ il valore ottenuto dall'equazione di Arrhenius 4.13 (0.0032 h^{-1}).

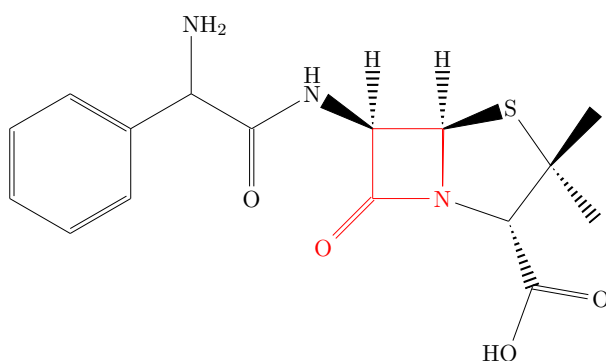


Figura 4.16: Struttura chimica dell'antibiotico ampicillina con evidenziato l'anello β -lattamico.

La colistina (paragrafo 3.4.1) è una polimixina formata da un eptapeptide legato a un tripeptide, che a sua volta lega un residuo di acido grasso all'estremità N terminale (Figura 4.17). Esiste in due isoforme, E_1 (colistina A) ed E_2 (colistina B), la cui differenza principale è rappresentata dall'acido grasso (FA), che è 6-MOA per E_1 e 6-MHA per E_2 .

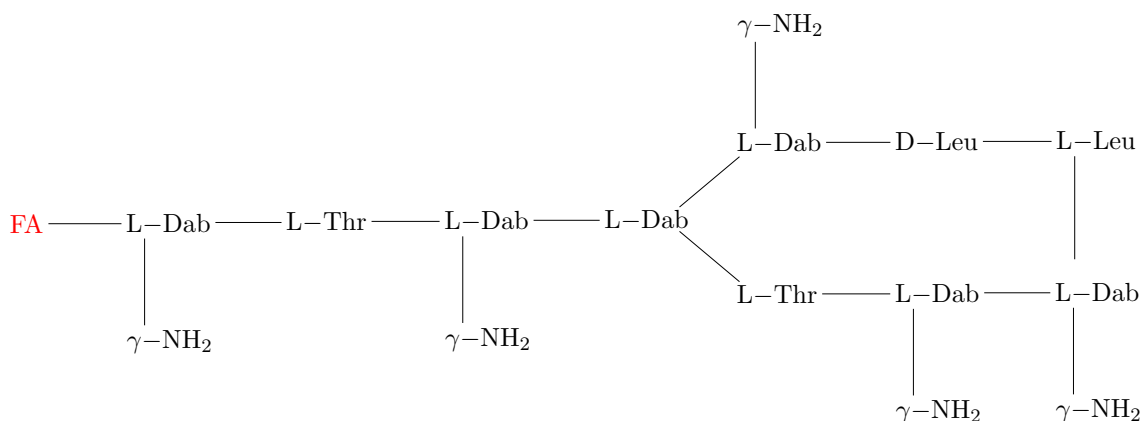


Figura 4.17: Struttura chimica dell'antibiotico colistina con evidenziato l'acido grasso FA diverso nelle due isoforme.

La colistina ha una stabilità maggiore in mezzi acidi e si degrada per racemizzazione con una cinetica di primo ordine dipendente da pH e temperatura. Orwa et al. [26] hanno

misurato i tassi di degradazione a pH 5.4 e 7.4 e a 37 °C in una soluzione di fosfati per entrambe le isoforme e hanno ottenuto i valori presenti in tabella 4.6.

	E ₁	E ₂	pH
k_1	$0.36 \cdot 10^{-3}$	$0.30 \cdot 10^{-3}$	5.4
k_2	$10 \cdot 10^{-3}$	$9.79 \cdot 10^{-3}$	7.4

Tabella 4.6: Tassi di degradazione delle isoforme E1 e E2 di colistina a temperatura 37 °C e pH 5.4 e 7.4.

Poiché la correlazione tra il logaritmo naturale di questi tassi e il pH a cui sono stati misurati è lineare, è possibile utilizzare un'interpolazione lineare per ottenere il tasso di degradazione γ a pH 7, che può essere inserito nel modello. L'equazione dell'interpolazione lineare è 4.14 e il tasso di degradazione ottenuto per la colistina E₁ vale 0.00514 h^{-1} , mentre quello per la colistina E₂ è 0.00488 h^{-1} . Il valore medio tra questi due tassi è 0.0050 h^{-1} , che è il valore del tasso di degradazione γ che si utilizza nel modello.

$$\ln(k) = \ln(k_1) + \frac{pH_7 - pH_{5.4}}{pH_{7.4} - pH_{5.4}} \cdot (\ln(k_2) - \ln(k_1)) \quad (4.14)$$

Karvanen et al. [27] hanno osservato che la colistina può essere persa durante gli esperimenti poiché aderisce alla plastica utilizzata in laboratorio essendo lipofila e idrofila. Per questo motivo, il valore reale di γ è maggiore rispetto a quello ottenuto. Poiché però il nostro modello è soltanto qualitativo e ci sono troppe variabili da considerare per ottenere il valore reale, come la presenza dei batteri e la forma delle piastre utilizzate, è possibile trascurare il contributo alla degradazione dovuto all'aderenza e utilizzare il valore 0.0050 h^{-1} come tasso γ nel modello.

Dai due tassi γ è possibile calcolare le emivite degli antibiotici con l'equazione 4.15. I due tempi sono 216 ore per l'ampicillina e 138 ore per la colistina. Poiché tutti gli esperimenti in presenza di ampicillina hanno una durata massima di 306 ore, questo antibiotico è sempre presente nelle colture, nonostante si riduca in concentrazione. Nel caso degli esperimenti con la colistina, invece, gli esperimenti più lunghi durano 455 ore, per cui l'antibiotico si degrada prima della fine dell'esperimento.

$$t_{1/2} = \frac{\ln(2)}{\gamma} \quad (4.15)$$

4.8 Studio del trasferimento del plasmide di antibiotico-resistenza

E' possibile seguire il trasferimento dei plasmidi di antibiotico-resistenza attraverso l'utilizzo di piastre selettive, chiamate "AMP" per il plasmide di resistenza all'ampicillina e "COL" per quello di resistenza alla colistina. Queste piastre hanno una concentrazione di antibiotico tale da far crescere soltanto i ceppi resistenti e da inibire quelli sensibili. Attraverso il plate counting si possono contare le CFU e osservare in quale momento appaiono i ceppi resistenti. Il passaggio dei plasmidi può avvenire fin dall'inizio, tuttavia è necessario del tempo affinché vengano trasferiti e i ceppi inizino a sviluppare la resistenza, diventando visibili sulle piastre.

In tutti gli esperimenti, la comunità KAN possiede soltanto ceppi sensibili, per cui non avviene lo sviluppo di resistenza e non si formano colonie sulle piastre AMP e COL. Le uniche due eccezioni sono i due replicati della comunità KAN in 2xMIC di colistina del terzo esperimento (Figura 4.18), dove è stato notato lo sviluppo di resistenza a causa di una mutazione avvenuta nel genoma dei ceppi PM e PL e non per motivi di coniugazione.

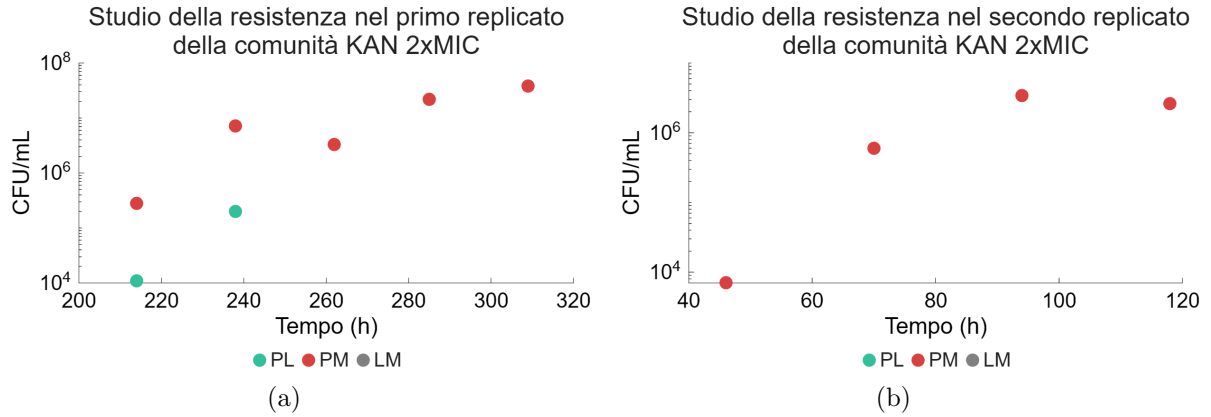


Figura 4.18: Crescita dei ceppi resistenti nei due replicati della comunità KAN con 2xMIC di colistina nelle piastre selettive "COL".

I secondi replicati delle comunità KAN + PVI, PVI e PVB del primo esperimento sono mostrati in figura 4.19 ed è possibile osservare che nella comunità KAN + PVI la resistenza è visibile dopo 68 ore, come si nota dall'aumento dei ceppi resistenti di PM e PL, mentre nella comunità PVI soltanto dopo 162 ore. Nella comunità PVB, invece, non avviene il passaggio del plasmide, per cui il ceppo LM tende a diminuire nel tempo, non potendo sopravvivere in mancanza degli altri due. Questo è in linea con le caratteristiche della comunità PVB.

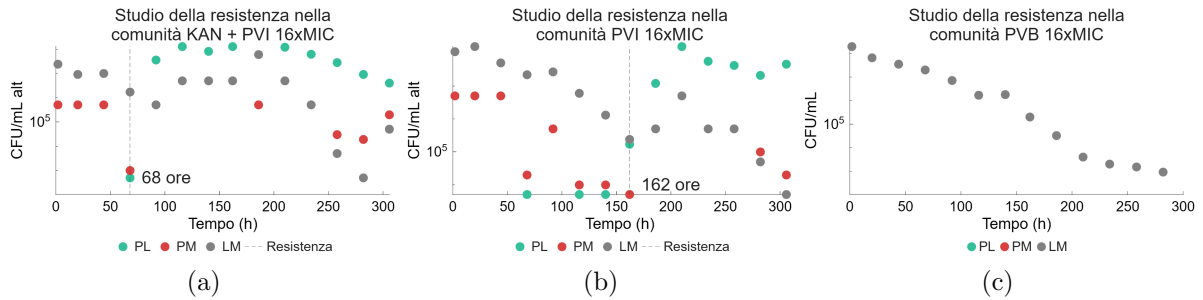


Figura 4.19: Crescita dei ceppi resistenti delle comunità KAN + PVI [4.19a], PVI [4.19b] e PVB [4.19c] nelle piastre selettive "AMP".

Si osservano in figura 4.20 i due replicati del terzo esperimento e quello del quarto per la comunità PN23 in presenza di una concentrazione 1xMIC di colistina. Anche in questo caso, emerge che la resistenza diventa visibile a tempi diversi: nel primo replicato del terzo esperimento appare dopo 94 ore, nel secondo replicato dopo 190 ore e nel quarto esperimento dopo 116 ore.

Infine, si possono studiare le comunità PN23 e KAN + PN23 in presenza di 2xMIC di colistina (figura 4.21). In questo caso, si nota che la popolazione PN23 diventa resistente soltanto nel primo replicato e si mostra dopo 214 ore. Gli altri due replicati non mostrano il passaggio del plasmide e questo può essere dovuto alla concentrazione di antibiotico troppo

4 Risultati e discussione

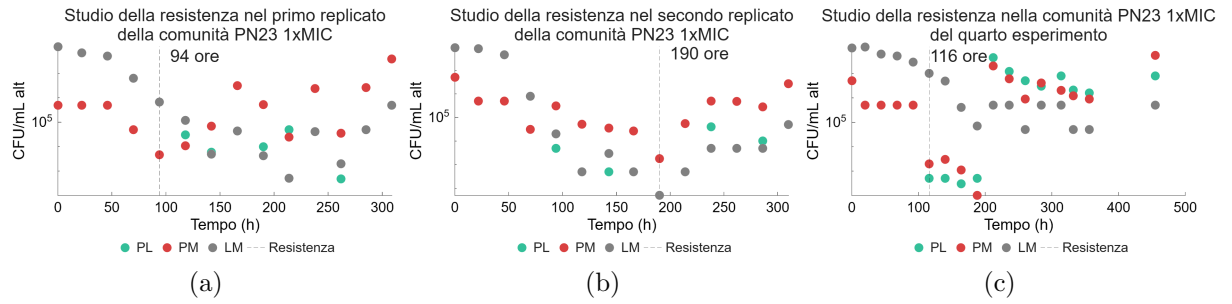


Figura 4.20: Crescita dei ceppi resistenti della comunità PN23 1xMIC nel primo [4.20a] e secondo [4.20b] replicato del terzo esperimento e nel quarto esperimento [4.20c] nelle piastre selettive "COL".

elevata, che non permette alle cellule di sopravvivere abbastanza a lungo da sviluppare la resistenza. La comunità KAN + PN23, invece, riesce a sviluppare la resistenza nonostante l'alta concentrazione antibiotica, e mostra i ceppi resistenti dopo 190 e 46 ore.

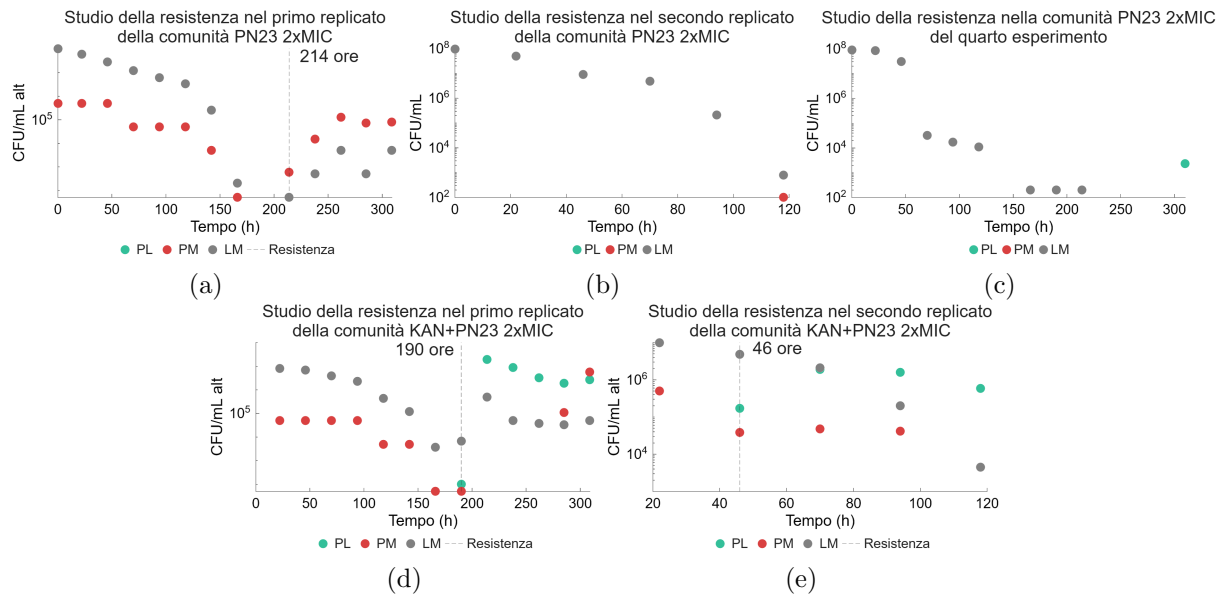


Figura 4.21: Crescita dei ceppi resistenti della comunità PN23 e KAN + PN23 2xMIC nel primo [4.21a, 4.21d] e secondo [4.21b, 4.21e] replicato del terzo esperimento e nel quarto esperimento [4.21c] nelle piastre selettive "COL".

In conclusione, possiamo osservare che la resistenza nelle diverse comunità diventa visibile a tempi diversi nei singoli esperimenti, non permettendo di estrapolare una regola generale sui tempi impiegati per la coniugazione e l'espressione dei plasmidi di resistenza. Si nota inoltre che per concentrazioni troppo elevate di antibiotico la coniugazione non riesce ad avvenire, poiché le popolazioni sensibili muoiono prima di diventare resistenti. Infine, si può affermare che la comunità PVB ha dato i risultati attesi, non riuscendo a trasferire il plasmide.

4.9 Costruzione del modello completo

E' possibile costruire un modello unico che rappresenti tutte le comunità in presenza e in assenza di antibiotico (script di MATLAB presente nei Materiali Supplementari [6.7]). Sono utilizzate le equazioni differenziali 4.9, che presentano i termini legati alla presenza dell'antibiotico. Queste equazioni possono rappresentare anche le comunità in assenza di antibiotico azzerando la concentrazione di antibiotico Ab , il tasso di morte dovuto all'antibiotico (ε), il tasso di coniugazione (η) e i valori dei ceppi resistenti B_R e C_R . Poiché si utilizzano come parametri quelli ottenuti dal fitting di tutte le comunità senza antibiotico del primo esperimento, è possibile eliminare anche il termine A_R e considerare soltanto A_S come rappresentativo di tutta la comunità. Con queste equazioni differenziali e i parametri ottenuti dal fitting sui dati, è possibile simulare le comunità in presenza di ampicillina oppure colistina e osservare come cambia il loro andamento al variare dei parametri ε e η . E' possibile anche confrontarlo con i dati sperimentali ottenuti dal primo esperimento e osservare se il modello riesce a simulare l'andamento oppure no. Poiché nelle comunità con ampicillina è presente un solo esperimento per tipo e nelle comunità con colistina si hanno comportamenti diversi in ogni esperimento, i dati sperimentali sono utilizzati soltanto qualitativamente per dare l'idea di come dovrebbe essere l'andamento. E' possibile osservare graficamente anche come variano le concentrazioni dei ceppi resistenti B e C in presenza dell'antibiotico. Normalmente, all'inizio la popolazione è formata soltanto da ceppi sensibili di B e C, per poi trasformarsi con il tempo in resistenti attraverso la coniugazione del plasmide. Un esempio di simulazione della comunità PVI in presenza di ampicillina è visibile in figura 4.22, in cui i parametri ε ed η valgono rispettivamente 0.1 e 0.0005 e il tasso di crescita dei ceppi resistenti è 1.1 volte quello dei sensibili.

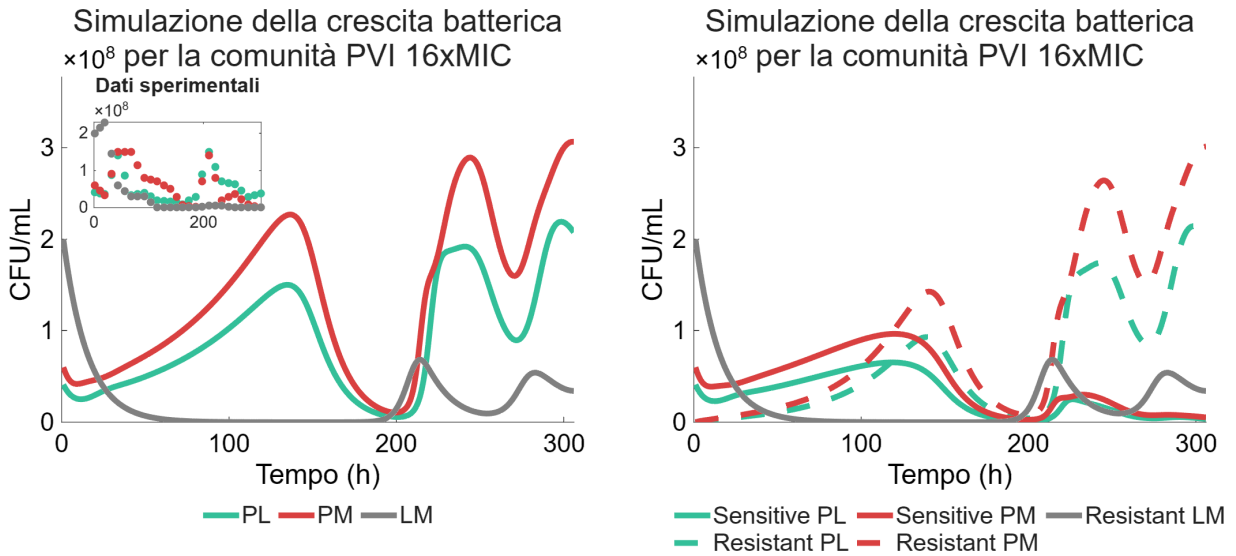


Figura 4.22: A sinistra, simulazione dell'andamento di LM, PM e PL per la comunità PVI e confronto con i dati sperimentali. A destra, suddivisione dei ceppi PM e PL in resistenti e sensibili.

4.10 Studio del cambiamento dei parametri ε e η tramite grid map e heatmap

E' possibile studiare come cambia l'andamento delle simulazioni e l'abbondanza dei ceppi sensibili e resistenti nella comunità PVI al variare dei parametri ε e η attraverso due tipi di grafici di MATLAB: grid map e heatmap. Il primo grafico permette di osservare le curve dei ceppi sensibili e resistenti e di vedere come si modificano in base ai parametri; non è possibile, tuttavia, osservare un numero troppo elevato di combinazioni di ε e η . Le heatmap, invece, si possono utilizzare anche per un numero molto elevato di combinazioni dei parametri e rappresentano come cambiano le simulazioni attraverso una scala di colori.

Attraverso la grid map in figura 4.23, si osservano i grafici che si formano dalla combinazione di dieci valori di ε tra 0.1 e 1 e di dieci valori di η tra 0.00001 e 0.0004. Questi valori sono stati scelti per tentativi identificando quelli che mostrano maggiori variazioni nelle curve. Quello che si può dedurre è che all'aumentare di ε , poiché la concentrazione di antibiotico rimane costante, i ceppi oscillano **meno** e il loro picco si sposta a tempi maggiori. Osservando invece la variazione di η , si nota che inizialmente questo parametro è troppo basso e non porta alla nascita di ceppi resistenti, mentre successivamente questi aumentano fino a prevalere sui sensibili. Quando si ha una combinazione di ε ed η a valori alti, si osserva che rimangono soltanto i ceppi resistenti, poiché i sensibili spariscono a causa della morte indotta dall'antibiotico e della trasformazione rapida in resistenti. E' emerso anche che i valori di η superiori a 0.0004 portano un cambiamento trascurabile delle abbondanze nei ceppi, producendo dei grafici molto simili a quello osservato per $\eta = 0.0004$.

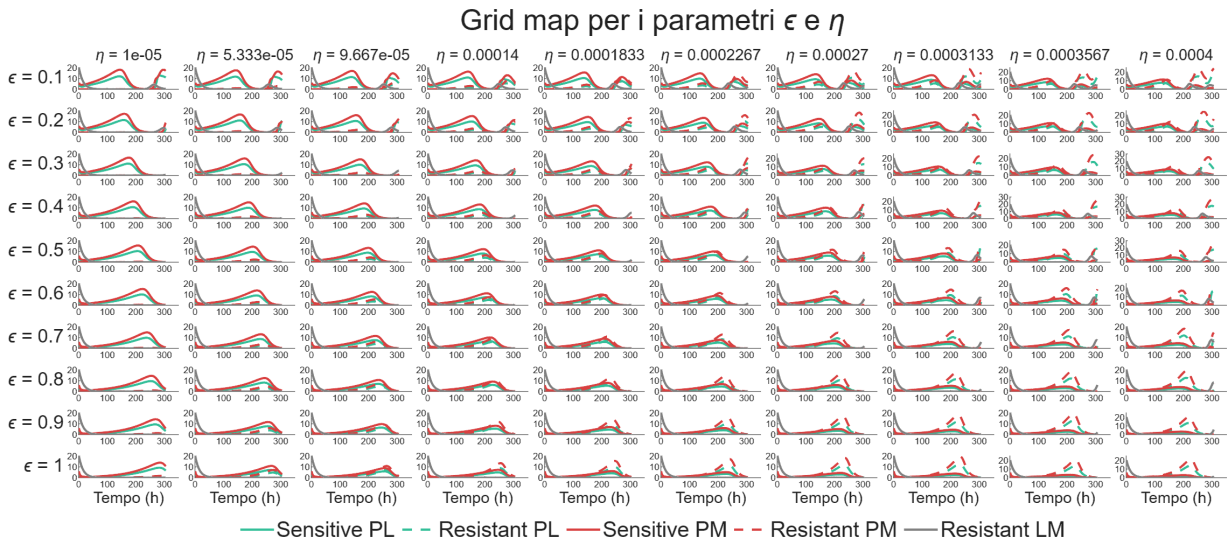


Figura 4.23: Grid map della comunità PVI al variare dei parametri ε ed η .

La heatmap è un tipo di grafico che permette di osservare cosa accade a una variabile al cambiamento di altre due. In questo caso, si può utilizzare per osservare come cambia il tempo di acquisizione della resistenza al variare di ε ed η . Si considerano 25 valori di ε tra 0.1 e 1, che aumentano da sinistra a destra, e 25 valori di η tra 0.0005 e 0.0055, che aumentano dall'alto in basso. Poiché le comunità iniziano a diventare resistenti fin dalla prima ora, si è scelto di osservare i tempi in cui i ceppi B e C resistenti superano 3 CFU. In figura 4.24, è possibile osservare nella prima riga le heatmap dei tempi in cui

B_r e C_r superano 3 CFU, mostrando sia il caso in cui entrambi i ceppi superano la soglia contemporaneamente, sia i casi in cui la superano singolarmente. E' possibile notare che per alti valori di η e bassi valori di ε i ceppi superano molto rapidamente la soglia di 3 CFU, mentre al diminuire di η e all'aumentare di ε si trovano i tempi maggiori. Inoltre, il ceppo B diventa resistente più velocemente rispetto al ceppo C, che è quindi il ceppo limitante tra i due.

Nella seconda riga della figura 4.24, sono stati osservati i tempi in cui metà delle popolazioni di B e C diventa resistente per avere un'idea della composizione della comunità al variare di ε ed η . Come prevedibile, i tempi sono minori per alti valori di η e ε , mentre sono alti quando il tasso di coniugazione e il tasso di morte indotta dall'antibiotico sono piccoli, poiché la popolazione rimane sensibile più a lungo. Inoltre, le heatmap appaiono quasi uguali, infatti la variazione dei tempi tra loro è minima.

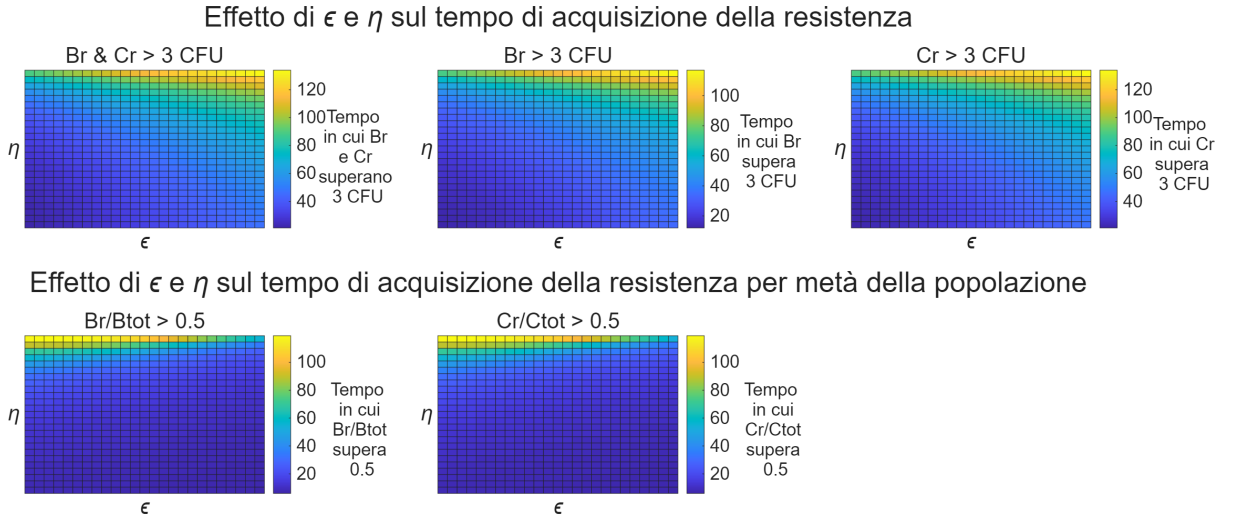


Figura 4.24: Heatmap degli effetti di ε ed η sui tempi di acquisizione della resistenza nella comunità PVI.

Le heatmap sono utilizzabili anche per studiare l'effetto della variazione dei tassi di produzione degli aminoacidi sui tempi di acquisizione della resistenza al variare di ε e η . Sono stati osservati i tempi in cui sia B che C resistenti superano 3 CFU e in cui metà della popolazione di B e C acquisisce la resistenza. I valori di ε ed η utilizzati sono gli stessi della heatmap precedente, mentre i valori di α sono stati modificati in due modi.

Nel primo caso (figura 4.25), sono stati utilizzati i valori di α corrispondenti agli estremi superiore e inferiore dell'intervallo di confidenza al 95%, oltre ai valori di α ottimizzati nel fitting come termini di confronto. E' possibile notare che, nel caso del valore dell'intervallo inferiore (LIC α), le simulazioni con ε alto e η basso non superano mai la soglia delle 3 CFU resistenti. Inoltre, i tempi in cui viene superata la soglia diminuiscono all'aumentare di α , passando da un massimo di 300 (LIC α) a 120 (α fittato) a circa 70 (UIC α). E' possibile osservare lo stesso andamento, anche se meno marcato, nel tempo di acquisizione della resistenza di metà delle popolazioni B e C, in cui i tempi passano da un massimo di circa 160 (LIC α) a un massimo di 110 (α fittato) a un massimo di 90 (UIC α). Anche in questo caso, la differenza tra le heatmap dei tempi di resistenza di metà popolazione di B e di C è minima a parità di valori di α utilizzati.

Nel secondo caso, mostrato in figura 4.26, i valori di α ottenuti dal fitting sono stati diminuiti della metà e aumentati di 1.5, per osservare il cambiamento dei tempi di acquisizione della resistenza quando i parametri sono variati molto. Si nota in particolare che con i

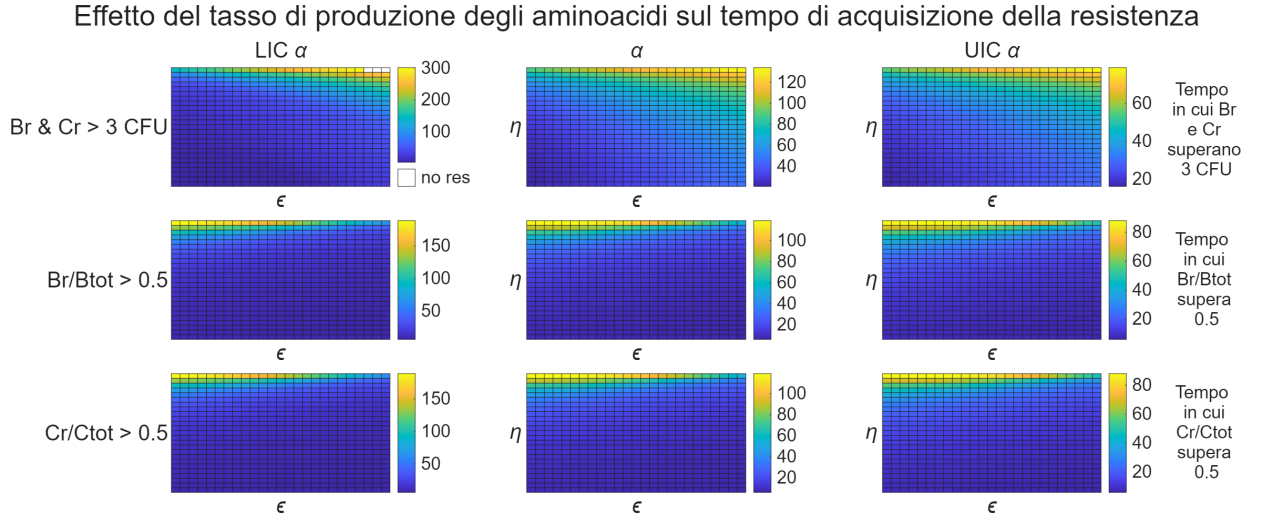


Figura 4.25: Heatmap degli effetti della variazione di α sui tempi di acquisizione della resistenza nella comunità PVI al variare di ϵ ed η . I valori di α utilizzati sono quelli degli intervalli di confidenza.

tassi di produzione diminuiti della metà, nessuna delle simulazioni supera la soglia di 3 CFU, per cui questa è stata abbassata a 1 CFU. Nonostante questo, la maggior parte delle simulazioni non supera la soglia, tranne che nei casi di alto η e basso ϵ . Dalle heatmap che mostrano i tempi di acquisizione della resistenza in metà della popolazione è possibile vedere che le popolazioni diventano resistenti in pochissimo tempo, anche se non riescono a crescere abbastanza da superare la soglia delle 3 CFU poiché il loro tasso di produzione degli aminoacidi è troppo basso. Anche in questo caso, si osserva che i tempi in cui viene superata la soglia delle CFU diminuiscono passando dal valore fissato di α (con un massimo di 130 ore circa) a quello aumentato di 1.5 (con un massimo di circa 45 ore) e lo stesso andamento si osserva nei tempi in cui metà popolazione acquisisce la resistenza.

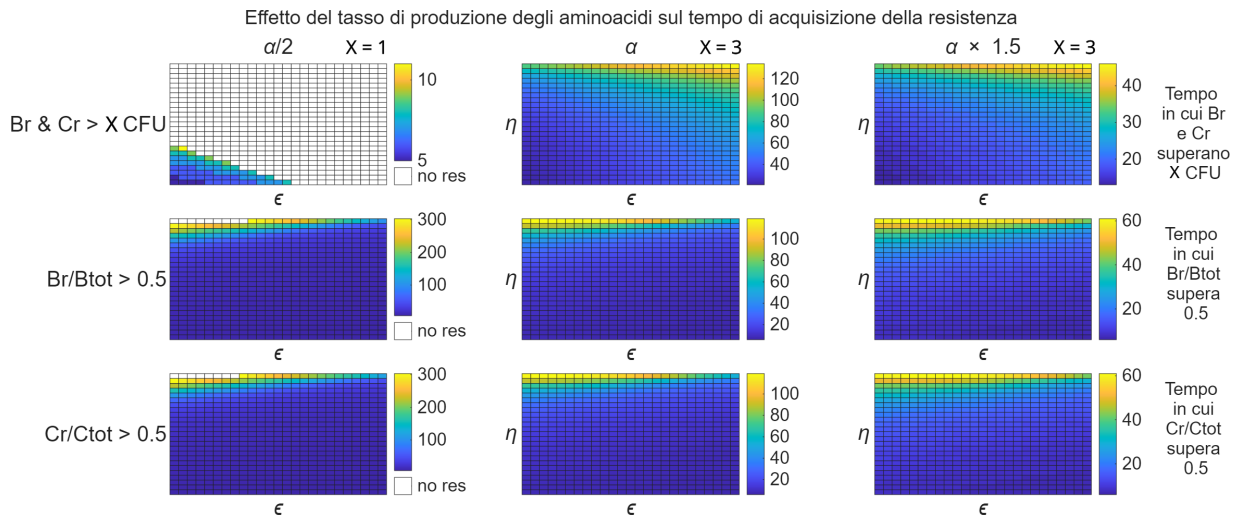


Figura 4.26: Heatmap degli effetti della variazione di α sui tempi di acquisizione della resistenza nella comunità PVI al variare di ϵ ed η . I valori di α utilizzati sono la metà e 1.5 volte quello ottenuto dal fitting.

I risultati delle heatmap con i valori di α più bassi sono molto diversi nonostante la

differenza tra i parametri non sia troppo elevata, come si può osservare nella tabella 4.7, che mostra come la differenza tra i valori dei parametri non sia elevata. Questo fa riflettere sul fatto che basta modificare di poco i tassi di crescita per passare da una popolazione resistente che cresce a una che non cresce, e l'impatto di α sulle popolazioni è molto alto.

Parametri	$\alpha/2$	LIC α	α	UIC α	$\alpha \times 1.5$
α_A	0.3823	0.5932	0.7645	0.9359	1.1468
α_B	0.0284	0.0471	0.0567	0.0663	0.0851
α_C	0.0601	0.0997	0.1202	0.1406	0.1803

Tabella 4.7: Valori di α_A , α_B e α_C utilizzati nelle heatmap di figura 4.25 e 4.26: valore fittato, estremi dell'intervallo di confidenza al 95% e valori dimezzato e moltiplicato per 1.5.

4.11 Difficoltà riscontrate durante la costruzione del modello

La costruzione del modello ha portato alla luce alcune problematiche legate ad aspetti sperimentali e modellistici. Ci sono cinque difficoltà principali collegate ai dati sperimentali.

La prima è legata al numero di replicati: per eseguire il fitting di un modello sui dati sperimentali e ottenere dei parametri affidabili, è necessario avere più replicati di una comunità, con la stessa durata e le misure sperimentali effettuate agli stessi tempi. Le comunità del modello che attualmente non hanno più di un replicato della stessa durata e negli stessi punti temporali sono KAN + PVI, PVI e PVB in presenza e in assenza di antibiotico e KAN, KAN + PN23 e PN23 in presenza di 2xMIC di antibiotico. KAN senza antibiotici possiede quattro replicati, di cui solo due della stessa durata e provenienti dal terzo esperimento. A causa tuttavia della diversità degli andamenti osservata nei due replicati, questi non possono essere utilizzati insieme. Per poter svolgere ugualmente il fitting sui dati sperimentali, si sono presi i tre replicati più simili della comunità KAN e si sono manipolati per avere la stessa lunghezza temporale e gli stessi tempi di ogni misura, come è stato spiegato in precedenza. L'alternativa migliore è di considerare tutti i secondi replicati delle quattro comunità senza ampicillina del primo esperimento, che hanno la stessa lunghezza e gli stessi punti temporali, e utilizzarli per modellare l'intera comunità senza antibiotico.

La seconda difficoltà è intrinseca nella rilevazione del numero di CFU del ceppo LM ed è visibile in tutte le comunità. Le colture del ceppo LM sono meno rilevabili rispetto a quelle dei ceppi PM e PL, per cui in molti punti sperimentali è risultato che la popolazione di LM fosse nulla, mentre le popolazioni di PM e PL continuavano a oscillare nel tempo. Questo è in contraddizione con la dinamica di interdipendenza dei tre ceppi, dal momento che la sopravvivenza delle specie può avvenire soltanto se sono presenti tutte e tre. La causa di questa contraddizione si trova nelle dimensioni inferiori di LM rispetto agli altri ceppi, che rendono la conta cellulare più complessa, e nel limite di rilevazione dell'1% dello strumento di plate counting, che non riesce a rilevare tutte le colonie di LM. Per superare questo problema, ci sono due metodi: il primo consiste nel manipolare i dati, stimando che i punti in cui LM risulta nullo in realtà abbiano la metà del limite di rilevazione delle

CFU; il secondo consiste nel cambiare la strumentazione, utilizzando un plate counting più sensibile.

La terza difficoltà riguarda il comprendere quali dei dati ottenuti sperimentalmente sono utili per il fitting del modello. I dati sperimentali sono la densità ottica, le fluorescenze, le CFU e gli eventi, tuttavia quelli da cui è possibile estrarre nel modo migliore l'andamento di crescita sono le CFU. E' possibile effettuare il fitting anche degli eventi, tuttavia le oscillazioni dei tre ceppi sono meglio osservabili dalle CFU. Il fitting sui dati di fluorescenza, invece, non permette di ottenere buoni risultati poiché i dati appartengono soltanto a due ceppi su tre e non è possibile modellare l'andamento del terzo (LM). Inoltre, è stato dimostrato che non è possibile estrapolare dai dati di fluorescenza e densità ottica il numero di CFU, per cui questi dati possono essere utilizzati solo qualitativamente per osservare gli andamenti, ma non possono essere fittati.

La quarta difficoltà si osserva negli andamenti dei dati dei quattro esperimenti. I replicati delle comunità sono molto diversi tra di loro, passando da oscillazioni multiple a oscillazioni uniche per lo stesso tipo di comunità. Questo fa pensare che i dati sperimentali siano molto variabili e che il fitting del modello sui dati delle comunità senza antibiotico non rispecchi l'andamento reale dei dati.

L'ultima difficoltà legata ai dati sperimentali si riferisce alla comunità KAN in presenza di 2xMIC di colistina. Nei due replicati del terzo esperimento è stato notato che la comunità diventa resistente a causa di una mutazione nel genoma, per cui questi replicati non sono realmente rappresentativi della comunità.

Le difficoltà legate al modello sono due principali. La prima è legata al fitting dei dati di LM, che presenta dei piccoli picchi di crescita anche quando i dati sperimentali non aumentano. Questo è dovuto al fatto che per simulare i picchi delle colture di PM e PL il modello abbia necessità di aumentare la presenza di LM e di conseguenza del suo aminoacido prodotto, indipendentemente dai dati sperimentali.

La seconda difficoltà riguarda l'andamento del modello a tempi successivi a quelli misurati sperimentalmente. Dal quarto esperimento sembrerebbe che la comunità KAN in assenza di antibiotici abbia un andamento ad oscillazioni smorzate nel tempo; questo in parte è visibile anche nella simulazione del modello con i parametri ottenuti dal fitting, tuttavia a tempi più lunghi le simulazioni crescono esponenzialmente in modo incompatibile con la dinamica reale della comunità.

5 Conclusioni

In questa tesi, è stato realizzato un modello di Lotka-Volterra generalizzato per studiare la dinamica di crescita delle comunità batteriche e la diffusione dell'antibiotico-resistenza attraverso il trasferimento genico orizzontale. Sono stati analizzati i dati sperimentali di densità ottica, fluorescenza, CFU ed eventi provenienti dai quattro esperimenti ed è stato studiato il comportamento delle singole comunità in assenza e in presenza di antibiotico, identificando i casi in cui riescono a sopravvivere e quelli in cui arrivano alla morte cellulare. Sono stati analizzati anche i tassi di crescita dei singoli ceppi su terreni arricchiti di aminoacidi e le conclusioni sono state due: ogni ceppo riesce a sopravvivere nel terreno in cui vengono forniti gli aminoacidi che non riesce a produrre; il plasmide PVI aumenta la fitness del ceppo LM. E' stata studiata la possibile conversione dei dati di fluorescenza e densità ottica in CFU, che non ha fornito i risultati sperati, non risultando possibile. Sono state definite le equazioni chimiche e differenziali del modello in assenza di antibiotico ed è stato eseguito il fitting del modello sui dati sperimentali provenienti dalle comunità KAN, KAN + PVI, PVI e PVB in assenza di ampicillina. Dal fitting, si sono ottenuti i parametri migliori da utilizzare nel modello ed è stata calcolata la loro affidabilità attraverso la misura degli intervalli di confidenza al 95%. E' stato utilizzato il modello per simulare l'andamento a tempi superiori a quelli sperimentali e la simulazione è stata confrontata con i dati del quarto esperimento, trovando una corrispondenza negli andamenti. Sono state definite successivamente le equazioni chimiche e differenziali del modello in presenza di antibiotico ed è stato studiato l'andamento della degradazione di ampicillina e colistina, dimostrando che l'ampicillina è presente lungo tutti gli esperimenti eseguiti, mentre la colistina, durante gli esperimenti più lunghi, potrebbe degradarsi e non essere più presente. Sono stati studiati i tempi in cui si iniziano a vedere i ceppi resistenti nelle varie comunità e sono stati ottenuti risultati molto variabili tra loro, che non hanno permesso di definire il tempo esatto in cui i ceppi diventano resistenti. Successivamente, è stato costruito il modello completo per simulare l'andamento di tutte le comunità in assenza e in presenza di antibiotici. Le simulazioni sono state confrontate con i dati sperimentali delle singole comunità per osservare le corrispondenze. Infine, è stato eseguito uno screening dei tassi di morte indotta dall'antibiotico (ε) e dei tassi di coniugazione del plasmide (η) attraverso grid map e heatmap, per osservare, al variare dei parametri, il cambiamento delle curve nelle simulazioni e dei tempi in cui le comunità superano una soglia di resistenza definita. Sono state prodotte anche heatmap per studiare il cambiamento dei tempi in cui viene acquisita la resistenza al variare dei tassi di produzione degli aminoacidi α_A , α_B e α_C .

5.1 Prospettive future e sviluppo del progetto

Come sviluppo futuro del progetto, potrebbero essere eseguiti nuovi esperimenti per ognuna delle comunità, in modo da ottenere più replicati della stessa durata e con misure effettuate negli stessi istanti temporali. Potrebbe poi essere svolto nuovamente il fitting sulla comunità KAN senza antibiotico, in modo da ottenere nuovi parametri che simulino in modo ancora più accurato l'andamento dei dati sperimentali. Allo stesso tempo, potrebbe

essere svolto un fitting anche sulle comunità con l'antibiotico per ottenere dei parametri che simulino l'andamento di ognuna delle comunità. Infine, potrebbero essere confrontati gli andamenti dei modelli, così da ottenere parametri simili per tutte le comunità e creare un modello universale, che possa simulare in modo fedele l'andamento di tutte le comunità e che possa fornire dati quantitativi.

Come sviluppo ulteriore, potrebbero essere svolti degli studi più dettagliati sulla resistenza alla colistina, cercando di identificare informazioni utili legate alla variazione dei parametri tramite heatmap o grid map.

Infine, uno sviluppo alternativo del progetto potrebbe consistere nell'aggiunta al ceppo LM di un plasmide che esprima una fluorescenza diversa da quelle di PM e PL. Questo permetterebbe di ottenere dati di fluorescenza su tutti i ceppi e potrebbe superare il problema dei dati da utilizzare nel fitting, rendendo il ceppo LM identificabile al plate counting oppure rendendo utilizzabili i dati di fluorescenza.

6 Materiale supplementare

6.1 Metodi alternativi di fitting

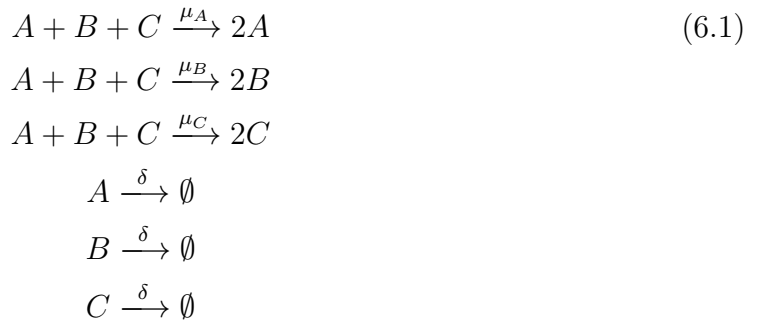
Durante la scelta dei dati sperimentali su cui eseguire il fitting del modello, sono state effettuate delle prove di fitting anche sugli eventi e sulle fluorescenze. Il fitting degli eventi si mostrava promettente, poiché aveva dati riferiti ai singoli ceppi, ed era un'alternativa alle CFU, poiché gli eventi di LM non raggiungevano mai un valore nullo. I problemi riscontrati durante il fitting, tuttavia, sono stati principalmente due e non hanno permesso l'utilizzo di questi dati. Il primo problema riguarda i parametri ottimizzati attraverso la funzione *SIMPLEXL* (paragrafo 3.2.1), tra cui il tasso di crescita di LM, che presentava sempre valori sotto 10^{-4} , apparendo come 0.0000. Il secondo problema riguarda la simulazione successiva ai tempi dei dati sperimentali, che cresce in modo esponenziale e senza oscillazioni, in modo incompatibile con l'andamento atteso degli eventi.

Il fitting delle fluorescenze, invece, non ha dato risultati utilizzabili, poiché i dati delle fluorescenze appartengono soltanto a due ceppi e il terzo non può essere fittato, per cui le simulazioni del modello non hanno mai seguito l'andamento dei dati sperimentali.

Tra i tentativi effettuati, è stato eseguito il fitting sui dati sperimentali provenienti dalle comunità KAN, KAN + PVI, PVI e PVB in presenza di ampicillina. I risultati ottenuti seguivano l'andamento dei singoli esperimenti, tuttavia i parametri ottimizzati erano diversi da quelli ottenuti per la comunità senza antibiotico e, essendoci un unico replicato per ogni comunità, non è stato possibile ottenere un fitting robusto sui dati. Un altro tentativo di fitting delle comunità con antibiotico consisteva nell'ottimizzazione solamente di ε , mantenendo fissi i parametri ottimizzati per la comunità senza antibiotico e utilizzando i dati sperimentali in presenza di antibiotico, che però non ha fornito i risultati sperati, non seguendo l'andamento sperimentale.

6.2 Modelli alternativi per la comunità senza antibiotico

E' stato ipotizzato un modello alternativo a quello utilizzato per le comunità senza antibiotico, al fine di eliminare le oscillazioni del fitting di LM non presenti nei dati. Questo modello voleva eliminare le equazioni riferite agli aminoacidi e sostituire, nelle equazioni rimanenti, il termine degli aminoacidi con il ceppo che li produceva. Le equazioni chimiche sono riportate in 6.1 e le equazioni differenziali derivate in 6.2.



$$\begin{aligned}
\frac{d[A]}{dt} &= \mu_A[A][B][C] - \delta[A] \\
\frac{d[B]}{dt} &= \mu_B[A][B][C] - \delta[B] \\
\frac{d[C]}{dt} &= \mu_C[A][B][C] - \delta[C]
\end{aligned} \tag{6.2}$$

Da questo modello non è stato possibile ottenere fitting che rappresentassero l'andamento di crescita dei ceppi batterici, poiché il modello non presentava oscillazioni.

6.3 Modelli per prevedere i dati

Quando è emerso che i dati del secondo esperimento non erano utilizzabili per eseguire il fitting e i dati del quarto esperimento non erano ancora disponibili, per ottenere delle serie più lunghe di dati è stato necessario attuare una strategia: sono stati aggiunti ai dati del primo esperimento dei dati sintetici oscillanti, che rispecchiavano l'andamento mostrato nel secondo esperimento. Per ottenere i nuovi dati, è stato utilizzato il modello ARIMA stagionale, che usa dati di serie temporali per predire l'andamento di nuovi dati a partire da quelli passati e ricalcare l'andamento delle oscillazioni. In mancanza di dati migliori, è possibile svolgere il fitting sui dati reali uniti a quelli sintetici, tuttavia non è detto che rispecchino l'andamento reale delle oscillazioni.

6.4 Modifiche della simulazione del modello

Una comunità batterica nel lungo periodo può raggiungere lo steady state, oscillare con oscillazioni smorzate oppure oscillare in modo persistente. Poiché dal secondo esperimento non si riusciva a comprendere in quale modo oscillasse la comunità e il modello simulava soltanto un andamento a oscillazioni smorzate, si volevano introdurre oscillazioni persistenti senza cambiare le equazioni differenziali. Per ottenere ciò, è possibile aggiungere del rumore alle concentrazioni delle specie batteriche nel modello, in modo che i ceppi continuino a ricevere uno stimolo e non raggiungano una concentrazione nulla. In questo modo, la comunità oscilla in modo persistente nel tempo, senza diminuire di ampiezza. Il rumore viene aggiunto soltanto alla seconda e alla terza specie, poiché la prima tende a zero fin dai dati sperimentali e la sua oscillazione non rispecchierebbe l'andamento ottenuto sperimentalmente.

6.5 Simulazione degli aminoacidi

Nella costruzione del modello si è studiata la dinamica di crescita della comunità, che è simulata dalle equazioni differenziali riferite ai ceppi. Dal modello stesso, tuttavia, è possibile simulare anche l'andamento degli aminoacidi (Figura 6.1), osservandone la variazione di concentrazione nel terreno di coltura. E' interessante osservare che l'andamento degli aminoacidi è in parte opposto a quello dei ceppi (Figura 4.14), infatti per esempio l'aminoacido prodotto da LM è in quantità molto alta, nonostante il ceppo raggiunga concentrazioni molto basse. Inoltre, gli aminoacidi, dopo alcune oscillazioni, raggiungono uno stato stazionario e non aumentano più nel tempo, a differenza di quanto osservato nelle simulazioni dei ceppi.

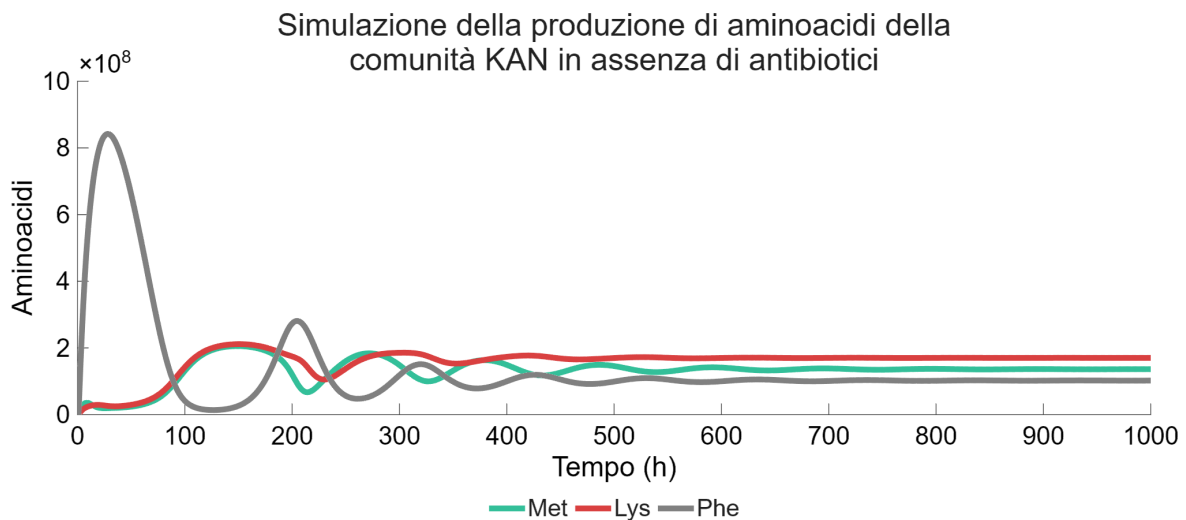


Figura 6.1: Simulazione dell'andamento degli aminoacidi per la comunità KAN senza antibiotici.

6.6 Script per calcolare gli intervalli di confidenza al 95%

Di seguito viene riportato lo script per calcolare gli intervalli di confidenza al 95% sui parametri ottimizzati con la funzione *SIMPLEXL* (paragrafo 3.2.1) a partire dai secondi replicati interpolati delle comunità KAN, KAN + PVI, PVI e PVB del primo esperimento.

```
%% Caricamento del file csv con le medie dei ceppi LM, PM e PL
% appartenenti alle comunità KAN, KAN + PVI, PVI e PVB senza
% antibiotico del primo esperimento.
cfu = load("../data/interpolated_mean.csv");
time_A = cfu(:,1);
cfu_A = cfu(:,2);
cfu_B = cfu(:,3);
cfu_C = cfu(:,4);
Data = [time_A, abs(cfu_A), abs(cfu_B), abs(cfu_C)];
```

```

%% Calcolo dei parametri migliori con la funzione SIMPLEXL
% Parametri iniziali:
p0 = [0.0004, 0.0005, 0.0005, 0.7759, 0.0520, 0.1324, 0.0702];

% Opzioni di SIMPLEXL
options = foptions;
options(1)=0;
options(2)=1e-8;
options(3)=1e-8;
options(14)=1000;

% Calcolo dei parametri migliori (parCal) attraverso SIMPLEXL e
% objectiveFunction
[parCal, ~] = SIMPLEXL('objectiveFunction', p0, options, [], Data);
parCal = parCal(:);
nParams = numel(parCal);

%% Vettore dei residui al punto ottimo
% Si chiede il secondo output di objectiveFunction, ossia il vettore
% dei residui
[~, r0] = objectiveFunction(parCal, Data);
n = numel(r0);

%% Jacobiano numerico dei residui rispetto ai parametri
% Lo Jacobiano è una matrice che contiene la derivata parziale del
% residuo i rispetto al parametro j ed esprime quanto varia ogni
% residuo al variare di ogni parametro.
J = zeros(n, nParams); % Matrice di 7 colonne e n righe
h = 1e-6; % dimensione del passo di perturbazione per stimare
% le derivate
for j = 1:nParams
    dp = zeros(nParams,1); % Vettore di perturbazione
    step = h * max(abs(parCal(j)), 1e-8); % E' la dimensione della
    % perturbazione e si moltiplica per il parametro per scalarla,
    % visto che i parametri hanno grandezze
    % diverse. 1e-8 serve se il parametro fosse 0, perché evita che
    % la derivata sia 0/0
    dp(j) = step; % Viene posizionato lo step nel vettore di
    % perturbazione

    % Misura objectiveFunction aumentando e diminuendo leggermente i
    % parametri
    [~, rPlus] = objectiveFunction(parCal + dp, Data);
    [~, rMinus] = objectiveFunction(parCal - dp, Data);

    % Differenza centrata:  $f'(x) = (f(x+step) - f(x-step)) / (2*step)$ 
    % Il suo errore è proporzionale a  $step^2$ 
    J(:,j) = (rPlus - rMinus) / (2*step);
end

```

```

% Utilizzo della funzione nlparci per calcolare gli intervalli di
% confidenza al 95%.
ci = nlparci(parCal, r0, 'Jacobian', J);
CI_lower = ci(:,1);
CI_upper = ci(:,2);

% Visualizzazione dei dati
paramNames = {'mu_A', 'mu_B', 'mu_C', 'alpha_A', 'alpha_B', 'alpha_C', ...
    'delta'};
fprintf('\n%-10s %12s %14s %14s\n', 'Parametro', 'Stima', ...
    'IC 95% low', 'IC 95% up');
for i = 1:nParams
    fprintf('%-10s %12.6g %14.6g %14.6g\n', ...
        paramNames{i}, parCal(i), CI_lower(i), CI_upper(i));
end

```

6.7 Script del modello completo

Di seguito, viene riportato lo script del modello completo, che permette di simulare le comunità in presenza e in assenza di antibiotico. E' composto da due file: il primo è nominato "TRICOMM" ed è lo script principale, che deve essere compilato per ottenere la simulazione; il secondo è chiamato "dYdt" e contiene la funzione al cui interno si trovano le equazioni differenziali del modello. Quest'ultimo viene richiamato nel primo script per svolgere le simulazioni. Può essere utilizzato per modellare sia la presenza di ampicillina che di colistina.

La concentrazione Ab_{init} di ampicillina è 0.18 poiché si ottiene dividendo la concentrazione 16xMIC (64 µg/ml) per il peso molecolare dell'ampicillina (349.4 g/mol); la concentrazione Ab_{init} di colistina, invece, è 0.00043 (1xMIC) o 0.00087 (2xMIC) e si ottiene dividendo la concentrazione 1xMIC (0.5 µg/ml) o 2xMIC (1 µg/ml) per il peso molecolare della colistina (1155.45 g/mol).

I tassi di crescita dei ceppi resistenti, invece, sono aumentati di un fattore 1.1, per indicare il miglioramento sul costo della fitness apportato da PVI. E' possibile integrare anche il costo sulla fitness di PVB e di PN23, una volta ottenuti i dati relativi al cambiamento che causano nel tasso di crescita batterico.

```

%% TRICOMM
%% Modello finale per la comunità Tricomm
% Questo file contiene il modello delle comunità con e senza
% antibiotico e usa i parametri derivati dalle comunità senza
% antibiotico.

% Lettura dei dati
cfu_AMP = readmatrix("../data/pOXA48_interpolated.csv");
cfu_COL = readmatrix("../data/PN23_interpolated.csv");

```

```

%% Scelta della comunità
% E' possibile scegliere quale antibiotico visualizzare, indicandolo
% nella variabile "antibiotic":
% 1 = ampicillina 16xMIC
% 2 = colistina

% Per la colistina, si può scegliere la "dose" per le comunità 1 e 3
% tra
% 1 = 1xMIC
% 2 = 2xMIC

antibiotic = 1;

if antibiotic == 2
    dose = 1;
end

% E' possibile scegliere anche quale comunità visualizzare, indicando
% il numero nella variabile "community":
% Per gli esperimenti con l'AMPICILLINA
% 0 = comunità senza antibiotico
% 1 = prima comunità con l'antibiotico (KAN) -> tutti i ceppi sono
% sensibili
% 2 = seconda comunità con l'antibiotico (KAN+PVI) -> tutti i ceppi
% sono sensibili e dopo due ore il ceppo LM resistente è aggiunto
% alla comunità ed è in grado di trasferire il suo plasmide
% 3 = terza comunità con l'antibiotico (PVI) -> LM è resistente e
% trasferisce il suo plasmide
% 4 = quarta comunità con l'antibiotico (PVB) -> LM è resistente e
% non trasferisce il suo plasmide

% Per gli esperimenti con COLISTINA
% 0 = comunità senza antibiotico -> stessi valori della comunità 0 di
% ampicillina
% 1 = comunità KAN con la colistina
% 2 = comunità KAN + PN23 in presenza di colistina
% 3 = comunità PN23 con la colistina

community = 3;

% Tutti i parametri delle comunità sono modificati automaticamente
% cambiando il numero della comunità. epsilon, eta e
% scale_factor possono essere modificati manualmente qui. epsilon
% è il tasso di morte indotta dall'antibiotico, eta è il tasso di
% coniugazione per il trasferimento del plasmide e scale_factor
% è il fattore di scala per il tasso di crescita dei ceppi
% resistenti, che vale 1 se è uguale a quello dei sensibili.

```

```

% Include il costo del plasmide sulla fitness batterica e qui è
% impostato a 1.1 perché è stato dimostrato che la presenza del
% plasmide PVI aumenta la fitness.

epsilon = 0.1;
eta = 0.0005;
scale_factor = 1.1;

%% Scelta dei dati sperimentali
% In questa parte, i dati sperimentali sono selezionati rispetto a
% quale comunità è osservata.

switch antibiotic
    case 1
        time = cfu_AMP(:,1);
    case 2
        time = cfu_COL(:,1);
end

switch community
    case 0 % comunità senza antibiotico
        cfu_A = cfu_AMP(:, 2);
        cfu_B = cfu_AMP(:, 3);
        cfu_C = cfu_AMP(:, 4);
    case 1 % KAN con l'antibiotico
        switch antibiotic
            case 1
                cfu_A = cfu_AMP(:, 5);
                cfu_B = cfu_AMP(:, 6);
                cfu_C = cfu_AMP(:, 7);
            case 2
                switch dose
                    case 1
                        cfu_A = cfu_COL(:, 2);
                        cfu_B = cfu_COL(:, 3);
                        cfu_C = cfu_COL(:, 4);
                    case 2
                        cfu_A = cfu_COL(:, 5);
                        cfu_B = cfu_COL(:, 6);
                        cfu_C = cfu_COL(:, 7);
                end
            end
        end
    case 2
        switch antibiotic
            case 1 % KAN + PVI con ampicillina
                cfu_A = cfu_AMP(:, 8);
                cfu_B = cfu_AMP(:, 9);
                cfu_C = cfu_AMP(:, 10);

```

```

        case 2 % KAN + PN23 con colistina
            cfu_A = cfu_COL(:, 8);
            cfu_B = cfu_COL(:, 9);
            cfu_C = cfu_COL(:, 10);
        end
    case 3
        switch antibiotic
            case 1 % PVI
                cfu_A = cfu_AMP(:, 11);
                cfu_B = cfu_AMP(:, 12);
                cfu_C = cfu_AMP(:, 13);
            case 2 % PN23
                switch dose
                    case 1
                        cfu_A = cfu_COL(:, 11);
                        cfu_B = cfu_COL(:, 12);
                        cfu_C = cfu_COL(:, 13);
                    case 2
                        cfu_A = cfu_COL(:, 14);
                        cfu_B = cfu_COL(:, 15);
                        cfu_C = cfu_COL(:, 16);
                    end
                end
            end
        case 4 % PVB
            cfu_A = cfu_AMP(:, 14);
            cfu_B = cfu_AMP(:, 15);
            cfu_C = cfu_AMP(:, 16);
        end

% Deviazione standard per le comunità senza antibiotico dagli
% esperimenti con ampicillina
std_A = cfu_AMP(:,17);
std_B = cfu_AMP(:,18);
std_C = cfu_AMP(:,19);

% Limite inferiore per la deviazione standard delle comunità senza
% antibiotico dagli esperimenti con ampicillina. Serve a imporre
% alla barra di errore inferiore il limite a zero se è negativa.
lower_A = cfu_AMP(:,20);
lower_B = cfu_AMP(:,21);
lower_C = cfu_AMP(:,22);

%% Parametri ottenuti dalle comunità senza antibiotico degli
% esperimenti con ampicillina.
mu_A = 0.0003;
mu_Ar = mu_A * scale_factor;
mu_B = 0.0005;
mu_Br = mu_B * scale_factor;
mu_C = 0.0004;

```

```

mu_Cr = mu_C * scale_factor;
alpha_A = 0.7645;
alpha_B = 0.0567;
alpha_C = 0.1202;
delta = 0.0692;
if community == 0 % Azzera i termini legati all'antibiotico
    epsilon = 0;
    eta = 0;
    Ab_init = 0;
end

if community == 1 || community == 4
    eta = 0; % Blocca il trasferimento del plasmide
end

params = [mu_A; mu_Ar; mu_B; mu_Br; mu_C; mu_Cr; alpha_A; alpha_B;
alpha_C; delta; epsilon; eta];

%% Simulazioni con ode45
% Condizioni iniziali
switch community
    case {0, 1}
        As_init = cfu_A(1);
        Ar_init = 0;
    case {2, 3, 4}
        As_init = 0;
        Ar_init = cfu_A(1);
end
Bs_init = cfu_B(1);
Br_init = 0;
Cs_init = cfu_C(1);
Cr_init = 0;
P_init = 0;
L_init = 0;
M_init = 0;
switch antibiotic
    case 1
        Ab_init = 0.18;
    case 2
        switch dose
            case 1
                Ab_init = 0.00043;
            case 2
                Ab_init = 0.00087;
        end
end
end
Y0 = [As_init; Ar_init; Bs_init; Br_init; Cs_init; Cr_init; P_init;
L_init; M_init; Ab_init];

```

```

[t, Y] = ode45(@(t,Y) dYdt(t,Y,params, community, antibiotic),
1:time(end), Y0);

% Risultato per ogni ceppo e aminoacido
As = Y(:, 1);
Ar = Y(:, 2);
Bs = Y(:, 3);
Br = Y(:, 4);
Cs = Y(:, 5);
Cr = Y(:, 6);
P = Y(:, 7);
L = Y(:, 8);
M = Y(:, 9);
Ab = Y(:, 10);

B = Bs + Br;
C = Cs + Cr;
A = As + Ar;

% Nomi delle comunità
if antibiotic == 1
    switch community
        case 0
            community_name = "senza antibiotico";
        case 1
            community_name = "KAN";
        case 2
            community_name = "KAN + PVI";
        case 3
            community_name = "PVI";
        case 4
            community_name = "PVB";
    end
else
    switch community
        case 0
            community_name = "senza antibiotico";
        case 1
            community_name = "KAN";
        case 2
            community_name = "KAN + PN23";
        case 3
            community_name = "PN23";
    end
end

switch antibiotic
    case 1
        antibiotic_name = "16xMIC";

```

```

case 2
    if community == 1 || community == 3
        switch dose
            case 1
                antibiotic_name = "1xMIC";
            case 2
                antibiotic_name = "2xMIC";
            end
        elseif community == 2
            antibiotic_name = "2xMIC";
        end
    end

end

%% Grafico con la somma di ceppi sensibili e resistenti
figure;
t1 = tiledlayout(1,2);
if community == 0
    nexttile;
    plot(t, C*1e7, 'Color', [0.20, 0.75, 0.60], 'LineWidth', 1.8, ...
        'DisplayName', 'Fit PL');
    hold on
    plot(t, B*1e7, 'Color', [0.85, 0.25, 0.25], 'LineWidth', 1.8, ...
        'DisplayName', 'Fit PM');
    plot(t, A*1e7, 'Color', [0.75, 0.75, 0.75], 'LineWidth', 1.8, ...
        'DisplayName', 'Fit LM');
    e1 = errorbar(time, cfu_C*1e7, lower_C*1e7, std_C*1e7, '.', ...
        'Color', [0.20, 0.75, 0.60], 'LineWidth', 0.85, ...
        'MarkerSize', 20, 'DisplayName', 'Exp. PL');
    e2 = errorbar(time, cfu_B*1e7, lower_B*1e7, std_B*1e7, '.', ...
        'Color', [0.85, 0.25, 0.25], 'LineWidth', 0.85, ...
        'MarkerSize', 20, 'DisplayName', 'Exp. PM');
    e3 = errorbar(time, cfu_A*1e7, lower_A*1e7, std_A*1e7, '.', ...
        'Color', [0.75, 0.75, 0.75], 'LineWidth', 0.85, ...
        'MarkerSize', 20, 'DisplayName', 'Exp. LM');
    e1.CapSize = 0;
    e2.CapSize = 0;
    e3.CapSize = 0;

    xlabel('Tempo (h)', 'FontSize', 15);
    ylabel('CFU/mL', 'FontSize', 15);
    legend('FontSize', 15, 'Box','off');
    t1 = title({'Fitting del modello per la crescita batterica', ...
        'della comunità senza antibiotico'}, 'FontWeight', 'normal');
    ax = gca; % get current axes
    ax.FontSize = 16;
    ax.TitleFontSizeMultiplier = 1;
    t1.FontSize = 19;
    ax.XColor = 'black';
    ax.YColor = 'black';

```

```

        box off
    else
        % Grafico principale
        ax_main = nexttile;
        hold(ax_main, 'on')

        plot(ax_main, t, C*1e7, 'Color', [0.20, 0.75, 0.60], ...
            'LineWidth', 4.8, 'DisplayName', 'PL'); hold on
        plot(ax_main, t, B*1e7, 'Color', [0.85, 0.25, 0.25], ...
            'LineWidth', 4.8, 'DisplayName', 'PM');
        plot(ax_main, t, A*1e7, 'Color', [0.75, 0.75, 0.75], ...
            'LineWidth', 4.8, 'DisplayName', 'LM');
        y_max_first200 = max([C(20:200); B(20:200); A(20:200)]) *1e7;
        % Serve a non sovrapporre il grafico principale con quello
        % piccolo
        hold(ax_main, 'off');
        xlabel(ax_main, 'Tempo (h)', 'FontSize', 20);
        ylabel(ax_main, 'CFU/mL', 'FontSize', 20);
        ylim(ax_main, [0, y_max_first200 + 15*1e7]);
        legend(ax_main, 'FontSize', 20, 'Box','off', 'Location', ...
            'southoutside', 'Orientation','horizontal');
        t2 = title({'Simulazione della crescita batterica', sprintf('per
            la comunità %s %s', community_name, antibiotic_name)}, ...
            'FontWeight', 'normal', 'FontSize', 19);
        ax_main.FontSize = 21;
        ax_main.TitleFontSizeMultiplier = 1;
        t2.FontSize = 26;
        ax_main.XColor = 'black';
        ax_main.YColor = 'black';
        box(ax_main, 'off');
    end

%% Grafico con i ceppi sensibili e resistenti separati

if community ~= 0
    nexttile;
    switch community
        case {1}
            plot(t, Cs*1e7, 'Color', [0.20, 0.75, 0.60], ...
                'LineWidth', 4.8, 'DisplayName', 'Sensitive PL');
            hold on
            plot(t, Bs*1e7, 'Color', [0.85, 0.25, 0.25], ...
                'LineWidth', 4.8, 'DisplayName', 'Sensitive PM');
            plot(t, As*1e7, 'Color', [0.75, 0.75, 0.75], ...
                'LineWidth', 4.8, 'DisplayName', 'Sensitive LM');
        case {2, 3}
            plot(t, Cs*1e7, 'Color', [0.20, 0.75, 0.60], ...
                'LineWidth', 4.8, 'DisplayName', 'Sensitive PL');
            hold on

```

```

        plot(t, Cr*1e7, '--', 'Color', [0.20, 0.75, 0.60], ...
            'LineWidth', 4.8, 'DisplayName', 'Resistant PL');
        plot(t, Bs*1e7, 'Color', [0.85, 0.25, 0.25], ...
            'LineWidth', 4.8, 'DisplayName', 'Sensitive PM');
        plot(t, Br*1e7, '--', 'Color', [0.85, 0.25, 0.25], ...
            'LineWidth', 4.8, 'DisplayName', 'Resistant PM');
        plot(t, Ar*1e7, 'Color', [0.75, 0.75, 0.75], ...
            'LineWidth', 4.8, 'DisplayName', 'Resistant LM');
    case 4
        plot(t, Cs*1e7, 'Color', [0.20, 0.75, 0.60], ...
            'LineWidth', 4.8, 'DisplayName', 'Sensitive PL');
        hold on
        plot(t, Bs*1e7, 'Color', [0.85, 0.25, 0.25], ...
            'LineWidth', 4.8, 'DisplayName', 'Sensitive PM');
        plot(t, Ar*1e7, 'Color', [0.75, 0.75, 0.75], ...
            'LineWidth', 4.8, 'DisplayName', 'Resistant LM');
    end
    xlabel('Tempo (h)', 'FontSize', 20);
    ylabel('CFU/mL', 'FontSize', 20);
    if community ~= 0 % For drug communities
        ylim([0, y_max_first200 + 15*1e7]);
    end
    legend('FontSize', 20, 'Box', 'off', 'Location', 'southoutside', ...
        'NumColumns', 3);
    t3 = title({'Simulazione della crescita batterica', sprintf('per
        la comunità %s %s', community_name, antibiotic_name)}, ...
        'FontWeight', 'normal', 'FontSize', 19);
    ax = gca; % get current axes
    ax.FontSize = 21;
    ax.TitleFontSizeMultiplier = 1; % Deactivate automatic scaling
    t3.FontSize = 26;
    ax.XColor = 'black';
    ax.YColor = 'black';
    box off
end

if community ~= 0
    % Grafico piccolo dentro quello principale
    pos = ax_main.Position;
    ax_inset = axes('Position', [pos(1)-0.02, pos(2)+0.30, ...
        0.38*pos(3), 0.3*pos(4)]);
    hold(ax_inset, 'on');

    plot(ax_inset, time, cfu_C*1e7, '.', 'Color', [0.20, 0.75, ...
        0.60], 'LineWidth', 1.6, 'MarkerSize', 20, 'DisplayName', ...
        'Exp. PL');
    plot(ax_inset, time, cfu_B*1e7, '.', 'Color', [0.85, 0.25, ...
        0.25], 'LineWidth', 1.6, 'MarkerSize', 20, 'DisplayName', ...
        'Exp. PM');

```

```

plot(ax_inset, time, cfu_A*1e7, '.', 'Color', [0.75, 0.75, ...
    0.75], 'LineWidth', 1.6, 'MarkerSize', 20, 'DisplayName', ...
    'Exp. LM');

hold(ax_inset, 'off');
title(ax_inset, {'Dati sperimentali', ''});
ax_inset.FontSize = 14;
ax_inset.Title.FontSize = 17;
ax_inset.Box = 'on';
end

%-----%
%% Simulazione dopo 306 ore per la comunità con l'antibiotico
% Il modello simula l'andamento oscillatorio fino a 1000 ore

if community == 0
    % 1:1000 indica che la simulazione arriva a 1000 ore.
    [t, Y] = ode45(@(t,Y) dYdt(t,Y,params, community, antibiotic), ...
        1:1000, Y0);

    % Risultato per ogni ceppo e aminoacido
    As = Y(:, 1);
    Ar = Y(:, 2);
    Bs = Y(:, 3);
    Br = Y(:, 4);
    Cs = Y(:, 5);
    Cr = Y(:, 6);
    P = Y(:, 7);
    L = Y(:, 8);
    M = Y(:, 9);
    Ab = Y(:, 10);

    % Grafico dei risultati
    nexttile;
    plot(t, Cs*1e7, 'Color', [0.20, 0.75, 0.60], 'LineWidth', 1.8,...
        'DisplayName', 'PL');
    hold on
    plot(t, Bs*1e7, 'Color', [0.85, 0.25, 0.25], 'LineWidth', 1.8,...
        'DisplayName', 'PM');
    plot(t, As*1e7, 'Color', [0.75, 0.75, 0.75], 'LineWidth', 1.8,...
        'DisplayName', 'LM');
    hold off
    xlabel('Tempo (h)', 'FontSize', 15);
    ylabel('CFU/mL', 'FontSize', 15);
    legend('FontSize', 15, 'Box', 'off');
    t1 = title({'Simulazione della crescita batterica delle', ...
        'comunità senza antibiotico'}, 'FontWeight', 'normal');
    ax = gca; % get current axes
    ax.FontSize = 16;

```

```

    ax.TitleFontSizeMultiplier = 1;    % Deactivate automatic scaling
    t1.FontSize = 19;
    ax.XColor = 'black';
    ax.YColor = 'black';
    box off
end

```

```

%% dYdt
function Ydot = dYdt(t, Y, params, community, antibiotic)
As = Y(1);
Ar = Y(2);
Bs = Y(3);
Br = Y(4);
Cs = Y(5);
Cr = Y(6);
P = Y(7);
L = Y(8);
M = Y(9);
Ab = Y(10);

mu_A = params(1);
mu_Ar = params(2);
mu_B = params(3);
mu_Br = params(4);
mu_C = params(5);
mu_Cr = params(6);
alpha_A = params(7);
alpha_B = params(8);
alpha_C = params(9);
delta = params(10);
epsilon = params(11);
eta = params(12);

switch antibiotic
    case 1
        gamma = 0.0032;
    case 2
        gamma = 0.0050;
end

switch community
    case {0,1}
        Ar = 0;
        Br = 0;
        Cr = 0;

```

```

    case {2,3}
        As = 0;
    case 4
        As = 0;
        Br = 0;
        Cr = 0;
    end

dAs_dt = mu_A * As * L * M - delta * As - epsilon * Ab * As;
dAr_dt = mu_Ar * Ar * L * M - delta * Ar;
dBs_dt = mu_B * Bs * P * M - delta * Bs - epsilon * Ab * Bs - eta ...
* Bs * (Ar + Br + Cr);
dBr_dt = mu_Br * Br * P * M - delta * Br + eta * Bs * (Ar + Br + Cr);
dCs_dt = mu_C * Cs * P * L - delta * Cs - epsilon * Ab * Cs - eta ...
* Cs * (Ar + Br + Cr);
dCr_dt = mu_Cr * Cr * P * L - delta * Cr + eta * Cs * (Ar + Br + Cr);
dP_dt = -mu_B * (Bs + Br) * P * M - mu_C * (Cs + Cr) * P * L + ...
alpha_A * (As + Ar);
dL_dt = -mu_A * (As + Ar) * L * M - mu_C * (Cs + Cr) * P * L + ...
alpha_B * (Bs + Br);
dM_dt = -mu_A * (As + Ar) * L * M - mu_B * (Bs + Br) * P * M + ...
alpha_C * (Cs + Cr);
dAb_dt = -(gamma + delta) * Ab;

Ydot = [dAs_dt; dAr_dt; dBs_dt; dBr_dt; dCs_dt; dCr_dt; dP_dt; dL_dt;
        dM_dt; dAb_dt];
end

```

Bibliografia

- [1] Naomi Iris van den Berg, Daniel Machado, Sophia Santos, Isabel Rocha, Jeremy Chacón, William Harcombe, Sara Mitri, and Kiran R Patil. Ecological modelling approaches for predicting emergent properties in microbial communities. *Nature ecology & evolution*, 6(7):855–865, 2022.
- [2] Ali R Zomorodi and Daniel Segrè. Synthetic ecology of microbes: mathematical models and applications. *Journal of molecular biology*, 428(5):837–861, 2016.
- [3] Katharine Z Coyte, Jonas Schluter, and Kevin R Foster. The ecology of the microbiome: networks, competition, and stability. *Science*, 350(6261):663–666, 2015.
- [4] Michael T Mee, James J Collins, George M Church, and Harris H Wang. Syntrophic **excme2014syntrophichange** in synthetic microbial communities. *Proceedings of the National Academy of Sciences*, 111(20):E2149–E2156, 2014.
- [5] Lisa Pagani, Ricardo León-Sampedro, Massimo Amicone, Burcu Tepekule, Christopher Witzany, Silvio D Brugger, Marjon GJ De Vos, Sara Mitri, Erik Bakkeren, Michael J Bottery, et al. Modelling the role of the microbiome in antimicrobial resistance across scales. *Nature Microbiology*, pages 1–14, 2026.
- [6] Tania Alonso-Vásquez, Michele Giovannini, Gian Luigi Garbini, Mikolaj Dziurzynski, Giovanni Bacci, Ester Coppini, Donatella Fibbi, and Marco Fondi. An ecological and stochastic perspective on persisters resuscitation. *Computational and Structural Biotechnology Journal*, 27:1–9, 2025.
- [7] Richard R Stein, Vanni Bucci, Nora C Toussaint, Charlie G Buffie, Gunnar Rätsch, Eric G Pamer, Chris Sander, and Joao B Xavier. Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota. *PLoS computational biology*, 9(12):e1003388, 2013.
- [8] Irene Creus-Martí, Andrés Moya, and Francisco J Santonja. CodaLomic: An r package for modeling microbiome compositional and longitudinal data. *PLOS Computational Biology*, 22(6):e1014328, 2026.
- [9] Hadley Wickham, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, Dewey Dunnington, Teun van den Brand, Posit, and PBC. Create elegant data visualisations using the grammar of graphics. <https://cran.r-project.org/web/packages/ggplot2/refman/ggplot2.html>. Publication: 2026-04-22 09:10:03 UTC.
- [10] Alboukadel Kassambara and Laszlo Erdey (Faculty of Economics and Business, University of Debrecen, Hungary). 'ggplot2' based publication ready plots. <https://cran.r-project.org/web/packages/ggpubr/refman/ggpubr.html>. Publication: 2026-02-24 14:30:02 UTC.

- [11] Hadley Wickham, Romain François, Lionel Henry, Kirill Müller, Davis Vaughan, Posit Software, and PBC. A grammar of data manipulation. <https://cran.r-project.org/web/packages/dplyr/refman/dplyr.html>. Publication: 2026-04-03 07:30:08 UTC.
- [12] Kevin Wright. Color palettes, colormaps, and tools to evaluate them. <https://cran.r-project.org/web/packages/pals/refman/pals.html>. Publication: 2025-03-07 18:20:02 UTC.
- [13] MATLAB. Curve fitting toolbox. <https://it.mathworks.com/help/releases/R2025b/curvefit/index.html>. Access: 2026-07-01.
- [14] MATLAB. Statistics and machine learning toolbox. <https://it.mathworks.com/help/stats/index.html>. Access: 2026-07-29.
- [15] MATLAB. ode45. <https://it.mathworks.com/help/matlab/ref/ode45.html>. Access: 2026-07-07.
- [16] MATLAB. nlparci. <https://it.mathworks.com/help/stats/nlparci.html>. Access: 2026-07-29.
- [17] MATLAB. Linear and nonlinear regression. <https://it.mathworks.com/help/releases/R2025b/curvefit/linear-and-nonlinear-regression.html>. Access: 2026-07-01.
- [18] MATLAB. Evaluating goodness of fit. <https://it.mathworks.com/help/curvefit/evaluating-goodness-of-fit.html>. Access: 2026-07-01.
- [19] Harrison Steel, Robert Habgood, Ciarán Kelly, and Antonis Papachristodoulou. Chibio: An open-source automated experimental platform for biological science research. *BioRxiv*, page 796516, 2019.
- [20] Kevin M Krause, Alisa W Serio, Timothy R Kane, and Lynn E Connolly. Amino-glycosides: an overview. *Cold Spring Harbor perspectives in medicine*, 6(6):a027029, 2016.
- [21] Joseph P Hou and John W Poole. Kinetics and mechanism of degradation of ampicillin in solution. *Journal of Pharmaceutical Sciences*, 58(4):447–454, 1969.
- [22] Carole Ayoub Moubareck. Polymyxins and bacterial membranes: a review of antibacterial activity and mechanisms of resistance. *Membranes*, 10(8):181, 2020.
- [23] Javier DelaFuente, Laura Toribio-Celestino, Alfonso Santos-Lopez, Ricardo León-Sampedro, Aida Alonso-del Valle, Coloma Costas, Marta Hernández-García, Lun Cui, Jerónimo Rodríguez-Beltrán, David Bikard, et al. Within-patient evolution of plasmid-mediated antimicrobial resistance. *Nature ecology & evolution*, 6(12):1980–1991, 2022.
- [24] Pramod K Jangir, Lois Ogunlana, Petra Szili, Marton Czikkely, Liam P Shaw, Emily J Stevens, Yang Yu, Qiue Yang, Yang Wang, Csaba Pál, et al. The evolution of colistin resistance increases bacterial resistance to host antimicrobial peptides and virulence. *Elife*, 12:e84395, 2023.

- [25] Shannon M Mitchell, Jeffrey L Ullman, Amy L Teel, and Richard J Watts. pH and temperature effects on the hydrolysis of three β -lactam antibiotics: Ampicillin, cefalotin and cefoxitin. *Science of the total environment*, 466:547–555, 2014.
- [26] JA Orwa, Cindy Govaerts, K Gevers, Eugene Roets, Ann Van Schepdael, and Jos Hoogmartens. Study of the stability of polymyxins b1, e1 and e2 in aqueous solution using liquid chromatography and mass spectrometry. *Journal of pharmaceutical and biomedical analysis*, 29(1-2):203–212, 2002.
- [27] Matti Karvanen, Christer Malmberg, Pernilla Lagerbäck, Lena E Friberg, and Otto Cars. Colistin is extensively lost during standard in vitro experimental conditions. *Antimicrobial agents and chemotherapy*, 61(11):10–1128, 2017.