

**DEFINIÇÃO DE VISÕES EM UM LIMS
ORIENTADO A FLUXO DE TRABALHO PARA
GERENCIAMENTO DE DADOS E PROCESSOS
LABORATORIAIS**

CRISTIANO SANTOS BOTELHO

**DEFINIÇÃO DE VISÕES EM UM LIMS
ORIENTADO A FLUXO DE TRABALHO PARA
GERENCIAMENTO DE DADOS E PROCESSOS
LABORATORIAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADOR: SÉRGIO VALE AGUIAR CAMPOS
COORIENTADOR: ALESSANDRA CONCEIÇÃO FARIA AGUIAR CAMPOS

Belo Horizonte

Agosto de 2012

© 2012, Cristiano Santos Botelho.
Todos os direitos reservados.

Santos Botelho, Cristiano

D1234p Definição de Visões em um LIMS Orientado a
Fluxo de Trabalho para Gerenciamento de Dados e
Processos Laboratoriais / Cristiano Santos Botelho.
— Belo Horizonte, 2012
xxv, 78 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de
Minas Gerais

Orientador: Sérgio Vale Aguiar Campos

1. LIMS. 2. Bioinformática. 3. Fluxo de Trabalho.
I. Título.

CDU 519.6*82.10



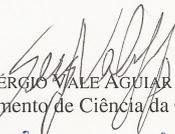
UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

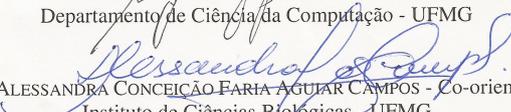
FOLHA DE APROVAÇÃO

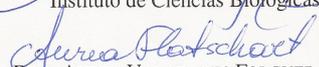
Definição de visões em um lms orientado a fluxo de trabalho para
gerenciamento de dados e processos laboratoriais

CRISTIANO SANTOS BOTELHO

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:


PROF. SÉRGIO VALE AGUIAR CAMPOS - Orientador
Departamento de Ciência da Computação - UFMG


DRA. ALESSANDRA CONCEIÇÃO FÁRIA AGUIAR CAMPOS - Co-orientadora
Instituto de Ciências Biológicas - UFMG


DRA. AUREA VALADARES FOLGUERAS -Flatschart
Instituto Nacional de Metrologia, Qualidade e Tecnologia - INMETRO


PROF. OMAR PARANAÍBA VILELA NETO
Departamento de Ciência da Computação - UFMG

Belo Horizonte, 03 de agosto de 2012.

Dedico esse trabalho a minha mãe Maria das Dores Ferreira Santos Botelho.

*“Existem mais coisas entre o céu e a terra
do que sonha a nossa vã filosofia.”*
(William Shakespeare)

Resumo

O uso da tecnologia da informação para automatizar os procedimentos em pesquisas na área da biomedicina, bioquímica e outras resulta em uma quantidade de dados muito grande e, por outro lado, a própria tecnologia da informação é que permite o gerenciamento dos dados gerados. Atualmente, o grande desafio é o armazenamento e gerenciamento desses dados de maneira eficiente e precisa. Assim, é comum o uso de Sistemas Gerenciadores de Informações de Laboratórios (LIMS) para cumprir essas tarefas. Os LIMS existentes são frequentemente restritos a uma determinada área. O SIGLa, por outro lado, por ser baseado em fluxo de trabalho. É capaz de atender laboratórios de áreas distintas. Os fluxos de trabalho são modelados na ferramenta *Together Workflow Editor* permitindo assim a flexibilidade do sistema. O Sistema SIGLa foi evoluído para outro sistema, Flux, que também é baseado em fluxo de trabalho. Esses sistemas podem incluir visões diferentes como parte de seu código. Visões são maneiras diferentes de se enxergar o conjunto de dados armazenados com o propósito de efetuar tarefas complementares às previstas na execução dos experimentos. O LIMS Flux, que é adaptável a diferentes laboratórios, foi adaptado com visões. Essas visões são: Controle de Insumos, Agendamento e Entradas e Saídas. A visão Controle de Insumos permite a definição de quatro tipos de fluxos de trabalho: cadastro, preparo e visualização de insumos e cadastro de receita. Com esses tipos de fluxos de trabalho é possível o gerenciamento de todos os insumos necessários para a realização de um experimento em laboratório. A visão Agendamento tem por objetivo permitir mecanismos de alerta e escalonamento de execução de atividades no tempo. E as Entradas e Saídas definem o uso de entradas e saídas atômicas, não-atômicas e o rastreamento dessas entradas e saídas. Como modelo de aplicação dessas visões, tem-se o LIMS *FluxTransgenics* que gerencia todos os dados e processos do laboratório de produção de transgênicos da EMBRAPA Milho e Sorgo (MG - Brasil).

Palavras-chave: LIMS, Fluxo de Trabalho, Bioinformática, Transgênicos.

Abstract

The use of information technology to automate process of biomedicine and biochemistry research and others result in a very large amount of data. Currently, the major challenge is the efficient and accurate storage and management of data. Thus, it is common to use Laboratory Information Management Systems (LIMS) to accomplish those tasks. Usually the available LIMS have its operations restricted to a specific area. In the other hand, SIGLa LIMS guarantees flexibility because it is based on workflows which are modeled in the Together Workflow Editor tool. The system SIGLa was upgraded to another system, Flux, which is workflow based too. These systems can aggregate different views as part of its code. Views are different forms of seeing the set of data stored in order to run complementary tasks to those provided in the experiment execution. The Flux LIMS that is adaptable to different laboratories was adapted with views. These views are Raw Material Control, Scheduling and Input and Output. The Raw Material Control view allows the definition of four kind of workflows: cadastre, prepare and visualization of raw material and cadastre of prescription. With these types of workflows it is possible the management of all raw materials needed for performing an experiment in laboratory. The Scheduling view aims to allow mechanisms of warning and scheduling the activity execution. And the Inputs and Outputs view defines the use of atomic and non-atomic inputs and outputs and their tracking. As model of application of these views, there is the FluxTransgenics LIMS that manages all data and process of the laboratory of production of Genetically Modified Organisms (OMS) at Embrapa Milho e Sorgo (MG-Brazil).

Keywords: LIMS, Workflow, Bioinformatics, Transgenics .

Lista de Figuras

| | | |
|-----|---|----|
| 2.1 | Elementos básicos de um fluxo de trabalho, processos, atividades e transição. | 12 |
| 3.1 | Exemplo de fluxo de trabalho com visões. O fluxo de trabalho <i>A</i> é o fluxo experimental enquanto os fluxos <i>B</i> e <i>C</i> são as visões de agendamento e controle de insumos, respectivamente. No fluxo <i>A</i> , os triângulos e os círculos coloridos em destaque representam subconjuntos de dados considerados visões do fluxo. | 21 |
| 3.2 | Exemplo de informação macro e específica relativas a um insumo. A informação macro é representada pela caixa e a informação específica por cada um dos frascos da mesma (<i>1, 2, 3, 4, 5</i> e <i>6</i>). | 22 |
| 3.3 | Elementos que compõem a visão de controle de insumos. As setas indicam a existência de um fluxo de informação de um componente para o outro. 1. Armazenamento do insumos cadastrados no banco de dados (BD); 2. Preparo de novos insumos com outros já cadastrados; 3. Acesso aos insumos no BD pelo fluxo de trabalho <i>Preparo de Insumos</i> ; 4. Relacionamento entre os componentes <i>Receita</i> e <i>Preparo de Insumo</i> ; 5. Inserção de novo insumo no BD; 6. Acesso dos insumos do banco de dados na execução de experimentos; 7. Visualização dos insumos que foram cadastrados ou preparados e registrados no banco de dados. | 24 |
| 3.4 | Construção do fluxo de trabalho Cadastro de Reagente no Together Workflow Editor. | 25 |
| 3.5 | Exemplo do <i>Cadastro de Receita</i> . O usuário insere um nome para identificar a receita, um arquivo em pdf que contem as instruções (protocolo) de preparo do <i>Acido Propionico</i> e a lista de insumos que serão necessários para preparar o insumo. | 26 |
| 3.6 | Resultado do uso da receita cadastrada na Figura 3.5. | 26 |

| | | |
|------|--|----|
| 3.7 | Construção do fluxo de trabalho Preparo de Insumo na ferramenta <i>Together Workflow Editor</i> . A definição do atributo <i>nomePop</i> se converterá no campo onde o usuário escolherá o protocolo para preparo da solução. . . . | 27 |
| 3.8 | Tela da visão Preparo de Insumos onde o usuário escolhe os insumos a serem usados no preparo do insumo. | 28 |
| 3.9 | Interface de visualização e edição dos insumos cadastrados e/ou preparados. Através do filtro o usuário pode fazer pesquisas variadas pelos insumos. . . | 28 |
| 3.10 | Exemplo de saída atômica na visão Entradas e Saídas | 31 |
| 3.11 | Interface da visão Entradas e Saídas para seleção das entradas a serem usadas na próxima atividade. | 32 |
| 3.12 | Exemplo de saída não atômica na visão Entradas e Saídas | 32 |
| 3.13 | Interface da visão Entradas e Saídas para escolha das entradas a serem usadas na atividade. | 33 |
| 3.14 | O fluxo de agendamento (X Agendamento) é gerado a partir do fluxo principal (X). | 34 |
| 4.1 | Tabela básica <i>workflow</i> , nessa tabela é gravado informações a cerca dos fluxos de trabalhos gerenciados pelo Flux. | 38 |
| 4.2 | Nessa tabela básica são armazenados os meta-dados das tabelas externas. . | 40 |
| 4.3 | Tabela básica onde são gravados os meta-dados das colunas das tabelas externas. | 41 |
| 4.4 | As informações das entradas e saídas são armazenadas nessa tabela básica. | 42 |
| 4.5 | Especificação de um fluxo de trabalho na ferramenta <i>Together Workflow Editor</i> sendo que o mesmo possui agendamento. | 43 |
| 4.6 | Tabela onde são armazenados os dados das atividades. | 44 |
| 4.7 | A tabela <i>schedule</i> faz parte do modelo de dados de tabelas básicas e armazena as informações referentes ao agendamento. | 45 |
| 5.1 | Abstração do relacionamento do fluxo de trabalho <i>Agrobacterium tumefaciens</i> com as visões Controle de Insumos, Agendamento e Entradas e Saídas. | 49 |
| 5.2 | Fluxo de trabalho da transformação de milho e sorgo via <i>Agrobacterim tumefaciens</i> . Cada caixa representa uma atividade do processo e as setas de uma atividade para a outra representam uma transição. | 50 |
| 5.3 | Tela do sistema <i>FluxTransgenics</i> demonstrando a primeira atividade do fluxo de trabalho da transformação via <i>Agrobacterium tumefaciens</i> . Nessa atividade, informações como a data, o responsável, o genótipo e a construção gênica são registrados. | 51 |

| | | |
|------|---|----|
| 5.4 | Tela do sistema <i>FluxTransgenics</i> que faz parte da atividade <i>Planejamento</i> . Esta tela é usada para escolher a construção gênica que será usada no experimento. | 52 |
| 5.5 | Tela do sistema <i>FluxTransgenics</i> para a atividade <i>Cocultivo de Embriões</i> . Após essa atividade ser executada, informações sobre a absorbância, quantidade e tamanho dos embriões, teste confirmatório, protocolos, etc são armazenados. | 53 |
| 5.6 | Tela do sistema <i>FluxTransgenics</i> para seleção de insumos na atividade <i>Cocultivo de Embriões</i> . Nesta atividade, após os embriões serem transformados, eles são colocados no meio de cultura feito a partir de uma solução preparada no laboratório. Essa tela é usada para escolher o insumo (solução) que constituirá o meio de cultura. | 53 |
| 5.7 | Tela do sistema <i>FluxTransgenics</i> mostrando entradas disponíveis para a atividade <i>Transferência para o Segundo Meio de Seleção</i> . Essas entradas são o resultado do cadastro das saídas na atividade anterior (<i>Transferência para o Primeiro Meio de Seleção</i>). | 54 |
| 5.8 | Tela do sistema <i>FluxTransgenics</i> mostrando cadastro de uma saída da atividade <i>Transferência para o Segundo meio de Seleção</i> no qual é possível escolher a entrada relacionada. | 55 |
| 5.9 | Tela do sistema <i>FluxTransgenics</i> de Cadastro de resultados (saída) da atividade <i>Transferência para o Segundo meio de Seleção</i> | 55 |
| 5.10 | Tela do sistema <i>FluxTransgenics</i> mostrando mensagem apresentada quando o fluxo de trabalho tem a visão de <i>Agendamento</i> e a atividade é executada fora da data agendada. | 56 |
| 5.11 | Fluxo de trabalho do processo de transformação de milho e sorgo via Bio-balística. Cada caixa representa uma atividade no processo e a seta de uma atividade para a outra representa uma transição. | 56 |
| 5.12 | Tela do sistema <i>FluxTransgenics</i> com a atividade <i>Bombardear Embriões</i> . Através dessa atividade, é possível preencher todas as informações a cerca do bombardeamento de embriões, tal como o número de tiros por placa e a pressão do gás helium no momento do bombardeamento dos embriões. . . . | 57 |
| 5.13 | Modelagem do fluxo de trabalho <i>Coleta de Embriões e Sub-cultivo de Calos</i> | 58 |
| 5.14 | Modelagem do fluxo de trabalho <i>Sub-cultivo de Controle Positivo</i> | 59 |
| 5.15 | Modelagem do fluxo de trabalho <i>Plantio de Sementes</i> | 60 |
| 5.16 | Tela do sistema <i>FluxTransgenics</i> para entrada da informação macro do cadastro de construção gênica. | 61 |

| | | |
|------|---|----|
| 5.17 | Tela do sistema <i>FluxTransgenics</i> para entrada das informações específicas que caracterizam cada frasco de construção gênica. | 61 |
| 5.18 | Tela do sistema <i>FluxTransgenics</i> que permite alterar os atributos específicos. | 62 |
| 5.19 | Tela do sistema <i>FluxTransgenics</i> para entrada da informação macro do cadastro de reagentes. | 63 |
| 5.20 | (A) Tela do sistema <i>FluxTransgenics</i> para detalhamento de informação específica. Na entrada das informações específicas, o campo medida é um <i>drop-down</i> menu com as opções peso, volume e n° de reações. A escolha de uma das opções resulta no aparecimento dos outros campos: peso inicial e atual se peso for escolhido; volume inicial e atual se volume for escolhido; e n° de reações inicial e atual caso o opção n° de reações seja selecionado. (B) Resultado da escolha do valor de medida peso. Ao escolher essa opção os campos Peso Inicial e Peso Atual foram habilitados para preenchimento. | 64 |
| 5.21 | Tela do sistema <i>FluxTransgenics</i> para Cadastro de Receita. Na informação macro do cadastro se insere o nome da receita e se faz o upload de um arquivo com o protocolo. Na parte de informações específicas, permite-se o registro dos nomes dos insumos usados na receita. | 65 |
| 5.22 | Tela do sistema <i>FluxTransgenics</i> demonstrando o uso da receita para preparo de solução. Os insumos relacionados à receita durante o cadastro da mesma são usados para gerar a tela onde o usuário seleciona os insumos usados no preparo. | 66 |
| 5.23 | Tela do sistema <i>FluxTransgenics</i> com formulário para entrada de dados do preparo de solução. | 67 |
| 5.24 | Tela do sistema <i>FluxTransgenics</i> do detalhamento de itens do fluxo de trabalho <i>Preparo de solução</i> | 67 |
| 5.25 | Tela do sistema <i>FluxTransgenics</i> para pesquisa de reagentes. 1: Filtro que permite a pesquisa dos insumos de acordo com um critério estabelecido. 2: Botão da interface para emissão de relatórios em formato pdf. 3: Botão da interface para emissão de relatórios em formato csv. | 68 |
| 5.26 | Tela do sistema <i>FluxTransgenics</i> com lista dos reagentes. A visualização fornece o total de cada reagente e o total global em estoque que compreende todos os reagentes. | 69 |

Lista de Tabelas

Lista de Abreviaturas e Siglas

BPL Boas Práticas de Laboratório

DNA Ácido Desoxirribonucleico

GDI Gene de Interesse

GLP Good Laboratory Practices

Kb Kilobase

LIMS Laboratory Information Management System

Pb Par de bases

TAT Transformação via *Agrobacterium tumefaciens*

TVB Transformação via Biobalística

TWE Together Workflow Editor

XML eXtensible Markup Language

XPDL XML Process Definition Language

WfMC Workflow Management Coalition

Sumário

| | |
|---|-------------|
| Resumo | xi |
| Abstract | xiii |
| Lista de Figuras | xv |
| Lista de Tabelas | xix |
| Lista de Abreviaturas e Siglas | xxi |
| 1 Introdução | 1 |
| 1.1 Motivação | 3 |
| 1.2 Contribuição | 4 |
| 1.3 Organização do texto | 4 |
| 2 Fundamentos | 7 |
| 2.1 Sistemas de Gerenciamento de Informação de Laboratórios (LIMS) . . | 7 |
| 2.1.1 Gerenciamento da explosão da informação | 7 |
| 2.1.2 Garantia da qualidade | 8 |
| 2.1.3 Redução de erros na entrada de dados | 9 |
| 2.1.4 Maior Eficiência e Eficácia na Execução das Atividades do laboratório | 10 |
| 2.2 Boas Práticas de Laboratório | 10 |
| 2.3 Fluxo de Trabalho | 11 |
| 2.3.1 XML Process Definition Language | 11 |
| 2.3.2 Together XPDL Workflow Editor | 14 |
| 2.4 Obtenção de Plantas Geneticamente Modificadas | 15 |
| 2.4.1 Transformação via <i>Agrobacterium tumefaciens</i> | 15 |
| 2.4.2 Transformação via Biobalística | 16 |

| | | |
|----------|---|-----------|
| 2.5 | Trabalhos Relacionados | 17 |
| 3 | Visões | 19 |
| 3.1 | Visão no Contexto de Banco de Dados | 19 |
| 3.2 | Visão no Contexto de LIMS | 19 |
| 3.3 | Visão de Controle de Insumos | 21 |
| 3.3.1 | Cadastro | 23 |
| 3.3.2 | Cadastro de Receita | 25 |
| 3.3.3 | Preparo de Insumos | 26 |
| 3.3.4 | Visualização | 28 |
| 3.3.5 | Utilização dos Insumos | 29 |
| 3.3.6 | Conclusão | 29 |
| 3.4 | Visão Entradas e Saídas | 30 |
| 3.4.1 | Saída Atômica | 30 |
| 3.4.2 | Saída Não-atômica | 31 |
| 3.4.3 | Rastreamento das Saídas | 31 |
| 3.4.4 | Conclusão | 32 |
| 3.5 | Visão de Agendamento | 33 |
| 4 | Implementação | 37 |
| 4.1 | Visão de Controle de Insumos | 38 |
| 4.2 | Entradas e Saídas | 39 |
| 4.3 | Agendamento | 40 |
| 5 | <i>FluxTransgenics</i> - Caso de Estudo | 47 |
| 5.1 | Fluxo de Trabalho da Transformação via <i>Agrobacterium tumefaciens</i> | 50 |
| 5.2 | Fluxo de Trabalho da Transformação via Biobalística | 55 |
| 5.3 | Coleta de Embriões e Sub-cultivo de Calos | 58 |
| 5.4 | Sub-cultivo de Controle Positivo | 59 |
| 5.5 | Plantio de Sementes | 59 |
| 5.6 | Controle de Insumos no <i>FluxTransgenics</i> | 60 |
| 5.6.1 | Cadastro de Construção Gênica | 60 |
| 5.6.2 | Cadastro de Reagente | 62 |
| 5.6.3 | Cadastro de Solução | 63 |
| 5.6.4 | Cadastro de Receita | 63 |
| 5.6.5 | Preparo de Solução | 66 |
| 5.6.6 | Visualização de Reagentes, Solução e Construção Gênica | 68 |
| 5.7 | Conclusão | 69 |

| | |
|---|-----------|
| 6 Conclusão | 71 |
| 6.1 Validação em Laboratório Real | 72 |
| 6.2 Trabalhos Futuros | 72 |
| Referências Bibliográficas | 75 |

Capítulo 1

Introdução

Pesquisas na área da biomedicina e bioquímica, em especial projetos de sequenciamento de DNA, tem gerado nos últimos anos desafios significativos no gerenciamento da informação, por parte dos laboratórios, uma vez que a quantidade de informação gerada é em larga escala. Adicionalmente, toda a pesquisa exige a utilização de um conjunto de regras para garantir a qualidade do trabalho, conhecidas como Boas Práticas de Laboratório (BPL), que exigem do pesquisador que os dados e resultados apresentados de uma pesquisa devem ser consistentes e rastreáveis. Uma das ferramentas para gerenciamento do grande volume de dados e a garantia do cumprimento das BPLs é o uso de Sistemas de Gerenciamento de Informações de Laboratório (*Laboratory Information Management Systems* - LIMS) para armazenar e gerenciar as informações geradas e permitir auditorias nos procedimentos executados em um laboratório.

LIMS são sistemas de gerenciamento da informação usados por laboratórios para gerenciar seus dados e os resultados de experimentos, os quais geralmente se encontram dispersos ou armazenados em sistemas não integrados. As principais razões para o uso de LIMS são o gerenciamento da explosão de informação, a garantia da qualidade, a redução de erros de entrada de dados e a garantia da eficiência e eficácia na execução de atividades em um laboratório [Hinton, 1995a]. Além disso, as características principais encontradas em um LIMS são a capacidade de armazenar e recuperar informações, interação com outros softwares, a manutenção de um plano de contingência, permitir auditorias, configurabilidade e flexibilidade e a apresentação de uma interface do usuário que suporte os princípios básicos da usabilidade (diálogo simples e natural, consistência, prevenção de erros, ajuda e documentação, boas mensagens de erro) [Hinton, 1995b].

Os LIMS existentes, como LabVantage Solutions [2010], Computing Solutions [2010] e Solutions [2010], são restritos a um tipo específico de domínio, e além disso são

pouco flexíveis. E não adaptáveis a outros tipos de laboratórios, por exemplo: um LIMS para gerenciamento de laboratório de análises clínica não atende a um de experimentos em sequenciamento de DNA. Sendo assim, o Laboratório de Universalização de Acesso (LUAR) desenvolveu um LIMS denominado SIGLa [Simões et al., 2010], adaptável a laboratórios de várias naturezas. Esse LIMS é baseado em fluxos de trabalho, o que permite a flexibilidade no suporte ao gerenciamento de diferentes Laboratórios. Porém, este não assegura algumas funcionalidades cruciais para adaptabilidade de um LIMS. Tendo o código fonte do SIGLa como ponto inicial, o Framework Flux foi desenvolvido com o objetivo de suprir funcionalidades ausentes no SIGLa. Assim como o SIGLa, o Flux é baseado em fluxo de trabalho.

Segundo a *WfMC (Workflow Management Coalition)* [WfMC, 2010], um fluxo de trabalho ou *workflow* é a automação do processo de negócio, na sua totalidade ou em partes, onde documentos, informações ou tarefas são passadas de um participante para o outro para execução de uma ação, de acordo com um conjunto de regras de procedimentos. A automação do processo de negócio identifica as várias atividades do processo, regras de procedimento e controle de dados associados para gerenciar o fluxo de trabalho durante a ativação do processo. Muitas instâncias individuais do processo podem se tornar operacionais durante a ativação do processo, cada uma associada a um conjunto específico de dados relevantes dessa instância.

O Flux foi construído como um LIMS que permite o gerenciamento de laboratórios com base na definição de fluxo de trabalho. Esses modelam os procedimentos do laboratório e são construídos em uma ferramenta de acesso livre, o *Together XPDL Workflow Editor (TWE)* [Editor, 2010]. Nessa ferramenta é possível a definição de fluxos de trabalho e se tem como saída toda a especificação em um arquivo XML seguindo o padrão XPDL. Dois fluxos de trabalho foram construídos inicialmente como exemplo de modelagem de processos no editor TWE para serem gerenciados pelo Flux: o fluxo de trabalho de análise proteômica e o de microarranjos de DNA. A proteômica compreende a identificação e análise quantitativa das proteínas codificadas pelos genes expressos nos diferentes tecidos ou estágios de desenvolvimento de um organismo [Bensmail & Haoudi, 2003; Panisko & Veenstra, 2002]. O estudo do conteúdo protéico de diversos tecidos e diferentes fases de desenvolvimento dos organismos é de extrema importância pois representa a chave para o entendimento de processos normais de desenvolvimento e o estabelecimento de patogenias. As aplicações práticas a partir de pesquisas com proteômica mais promissoras tem sido a projeção de novas drogas para tratamento de doenças [Abhilash, 2009]. Em laboratório de análise proteômica tem-se vários procedimentos tais como: a extração da proteína, cromatografia, eletroforese e a espectrometria de massa, dentre outros. Estes foram modelados no TWE no fluxo de

trabalho **Proteômica**. Já a técnica de microarranjos de DNA é utilizada para estudar a expressão gênica em grande quantidade em várias espécies, sendo que essa técnica permite o estudo de milhares de genes simultaneamente. Um microarranjo de DNA, ou DNA-chip, consiste num arranjo pré-definido de moléculas de DNA (fragmentos de DNA genômico, cDNAs ou oligo-nucleotídeos) quimicamente ligadas à uma superfície sólida, usualmente lâminas de microscópio revestidas com compostos que conferem carga positiva. Os microarranjos também podem ser preparados em membranas de nylon positivamente carregadas. As lâminas de microarranjo contendo sequências de nucleotídeos são utilizadas para duas finalidades principais: em estudo de genômica comparativa e na análise da expressão diferencial de genes [Draghici, 2003]. As atividades dessa técnica modeladas pelo editor TWE no fluxo de trabalho **Microarranjos** são: a extração de RNA, amplificação, marcação, hibridização, limpeza e concentração, isolamento, análise de imagem, análise de dados e validação dos dados.

Apesar de sua bem sucedida implementação no gerenciamento de dados biológicos, o sistema Flux apresenta ainda algumas limitações. O fato do fluxo de trabalho ser modelado apenas do ponto de vista da sucessão de atividades, torna a abstração de outras entidades relacionadas ao fluxo principal de trabalho inviável, uma vez que essas entidades são ortogonais ao fluxo de trabalho e são visões inerentes a cada fluxo de trabalho específico. As visões são formas distintas de enxergar os dados que não seguem o fluxo de trabalho experimental ou agrupamentos dos dados experimentais que compartilham propriedades em comum sendo que essas propriedades não estão diretamente relacionadas com o fluxo de trabalho. O LIMS Flux, que é adaptável a diferentes laboratórios, foi adaptado com visões. Essas visões são **Controle de Insumos**, **Agendamento** e **Entradas e Saídas**. Como exemplo de aplicação dessas visões, tem-se o LIMS *FluxTransgenics* que gerencia todos os dados e processos do laboratório de produção de organismos transgênicos do *Laboratório de Biologia Celular e Cultura de Tecidos* da EMBRAPA Milho e Sorgo (MG - Brasil).

1.1 Motivação

Os procedimentos relativos aos experimentos em um laboratório devem ser acompanhados e devidamente controlados para garantir a corretude dos dados obtidos e a execução de todas as etapas dos procedimentos. Para esse fim, são empregados LIMS, sendo que vários destes estão disponíveis no mercado. Contudo sua utilização é limitada devido ao alto custo e ao fato de que os LIMS existentes atualmente são específicos para determinados laboratórios. O sistema Flux é adaptável a diferentes laboratórios,

podendo suprir esta demanda. Entretanto, esse LIMS ainda não suporta o controle de algumas entidades inerentes ao fluxo de trabalho de um laboratório devido a características de sua modelagem. Assim é essencial agregar estas novas funções ao sistema afim de possibilitar a definição de visões particulares no Flux para cada tipo de fluxo de trabalho.

1.2 Contribuição

Esse trabalho propõe e implementa três visões para Sistemas de Gerenciamento de Informações de Laboratório (LIMS) baseado em fluxos de trabalho: **Controle de Insumos, Agendamento e Entradas e Saídas**. Visões são maneiras diferentes de se enxergar o conjunto de dados armazenados com o propósito de efetuar tarefas complementares às previstas na execução dos experimentos. Além disso, esse trabalho reporta a modelagem e desenvolvimento de fluxos de trabalho para gerenciamento do *Laboratório de Biologia Celular e Cultura de Tecidos* da Embrapa Milho e Sorgo que trabalha com a produção de sementes geneticamente modificadas. Os fluxos de trabalho foram desenvolvidos levando em consideração as visões propostas e implementadas.

1.3 Organização do texto

A fim de contribuir para a compreensão do trabalho, detalha-se nessa seção a estrutura do texto. No capítulo 1 foi apresentado o problema foco desse trabalho assim como a motivação para o mesmo e a contribuição do trabalho. O capítulo 2 apresenta alguns conceitos fundamentais para compreensão da leitura do texto, tais como Sistemas de Gerenciamento de Informação de Laboratórios (LIMS), processos de aquisição de plantas geneticamente modificadas (*Agrobacterium tumefaciens* e Biobalística) e faz um apanhado dos trabalhos relacionados. Nesse capítulo é detalhado o que são Boas Práticas de Laboratório (BPLs) e o que são fluxos de trabalho. O capítulo 3 apresenta o conceito de visões. No capítulo é dada a definição desse conceito no contexto de Sistemas Gerenciadores de Banco de Dados e em seguida a definição desse conceito no contexto desse trabalho. Ainda no capítulo 3 é detalhado cada visão definida no capítulo: Controle de Insumos, Entradas e Saídas e Agendamento. A fim de permitir uma visão geral dos detalhes técnicos relativos à implementação das visões, o capítulo 4 detalha as tecnologias usadas para desenvolver o trabalho e os algoritmos envolvidos. O capítulo 5 apresenta a modelagem do *Laboratório de Biologia Celular e Cultura de Tecidos* da Embrapa Milho e Sorgo. Nesse capítulo fica claro como as visões são

empregadas na construção de um LIMS baseado em fluxos de trabalho para gestão de dados e processos laboratoriais. O capítulo 6 apresenta as conclusões do trabalho, assim como possíveis extensões e melhorias.

Capítulo 2

Fundamentos

2.1 Sistemas de Gerenciamento de Informação de Laboratórios (LIMS)

Laboratory Information Management System, ou LIMS, são sistemas usados por laboratórios de diversas áreas para gerenciar e tornar acessível seus dados e resultados. Tais sistemas devem suportar o ciclo de vida dos dados o qual inclui a coleta, o armazenamento, a análise e a produção de relatórios dos mesmos. Existem várias razões para o emprego de LIMS em laboratórios, mas as seguintes razões afetam a maioria deles: gerenciamento da explosão da informação, garantia da qualidade, redução de erros na entrada de dados e maior eficiência e eficácia na execução das atividades do laboratório [Hinton, 1995a].

2.1.1 Gerenciamento da explosão da informação

Nos últimos anos, nota-se crescente explosão de informações nos processos de garantia da qualidade dos laboratórios. Isso acontece, principalmente por:

1. Determinações governamentais que exigem específicas restrições sobre as práticas e procedimentos de laboratório;
2. Avanços na automação que permitem rápida coleta de dados;
3. Elevados padrões de controle da qualidade em laboratórios;
4. Aumento da demanda por análises estatísticas dos dados, o que eleva a demanda por coleta, armazenamento, recuperação, análise e relatórios de dados.

Um LIMS permite a organização e armazenamento de dados de laboratório, contribui para garantia da qualidade dos procedimentos, a integração com outros departamentos, garante a capacidade de importação e exportação de outros pacotes de softwares e permite gerar relatórios e gráficos.

2.1.2 Garantia da qualidade

Laboratórios, cada vez mais, estão desenvolvendo programas de garantia da qualidade para cumprir com as regulamentações governamentais. As regulamentações podem estar relacionadas à produção de aditivos alimentares, à indústria farmacêutica e à análises de produtos químicos. Além dos órgãos públicos, existem organizações que trabalham para assegurar a garantia da qualidade, são exemplos:

- A *International Organization for Standardization (ISO)* que tem importante papel no comércio internacional. É a principal entidade responsável pela normalização e padronização em praticamente todos os campos técnicos em cerca de 170 países.
- Gerentes e pesquisadores em laboratórios seguem as práticas para garantia da qualidade de organizações como *APHA (American Public Health Association)*, *ASTM (American Society for Testing and Materials)*, *AOAC (Association of Official Analytical Chemists)*, *ACIL (American Council of Independent Laboratories)*, *ANSI (American National Standards Institute)* e *ASQC (American Society for Quality Control)*. Os objetivos de seguir as diretrizes dessas entidades são validar dados, aumentar a produtividade nos laboratórios e conseguir credibilidade no mercado.

Segundo Hinton [1995a] Controle de qualidade (CQ) é um sistema amplo de atividades, incluindo inspeção, que tem por objetivo prover a qualidade de um produto ou serviço que está regido perante um contrato. Na indústria, CQ significa atividades e técnicas usadas no dia-a-dia para verificar a qualidade do produto. É possível que os procedimentos para o CQ sejam substituídos por um método mais preciso de controle da qualidade. Hinton [1995a] apresenta as seguintes definições de qualidade:

- **Garantia da Qualidade:** consiste em todas as atividades necessárias e planejadas para prover um alto grau de confiabilidade na qualidade de um produto ou serviço. Além disso, avalia a qualidade das atividades de controle da qualidade e determina a validade dos procedimentos para determinação da qualidade.

- **Controle de Qualidade:** consiste nas técnicas e atividades necessárias para manter e melhorar a qualidade de um produto ou serviço.
- **Avaliação da Qualidade:** implica no conjunto de atividades que monitora e avalia a performance dos procedimentos de controle de qualidade usados na produção de um produto ou serviço.

LIMS automatizados podem contribuir substancialmente para a garantia da qualidade, haja vista que é mais produtivo e fácil gerenciar dados com sistemas computadorizados. Exemplos das contribuições que um LIMS pode fazer para a produtividade e garantia da qualidade são:

- Cálculos e entradas de dados são realizados em menor tempo;
- Melhora no tempo de buscas;
- Redução de erros de entrada de dados;
- Geração de relatórios de forma ágil;
- Capacidade de realização de auditorias;
- Automatização dos passos para realização de experimentos;
- Acesso distribuído aos dados do laboratório;
- Documentação das fases de produção.

2.1.3 Redução de erros na entrada de dados

Quando é discutida a garantia da qualidade, deve-se destacar o problema de erros na entrada de dados. Como a entrada de dados é umas das principais fontes de erros no contexto de validação de dados identificado por *EPA's Good Automated Laboratory Practices (GALPs)* Hinton [1995a], métodos para reduzir erros na entrada de dados devem ser examinados. A maioria dos LIMS tem especificidades para reduzir erros na entrada de dados. Alguns exemplos são:

- Restrições na entrada de dados;
- Verificação de intervalo;
- Código de barras;
- Entradas na forma de lista (*Drop-down menu*);

- Cálculos automáticos;
- Geração automáticas de relatórios;
- Validação dos dados de acordo com algum critério;
- Boas Práticas de Laboratório para validação de dados durante a entrada de dados.

2.1.4 Maior Eficiência e Eficácia na Execução das Atividades do laboratório

Eficiência se refere à relação entre os resultados obtidos e os recursos empregados. É a capacidade de conseguir produtos mais elevados em relação aos recursos necessários para obtê-los. Já a eficácia se refere a relação entre os resultados obtidos e os objetivos pretendidos, ser eficaz é atingir um dado objetivo. Eficiência e eficácia podem ter longo ou curto efeito. Por exemplo, eficiência e eficácia na medicina podem ajudar um paciente a receber atenção médica mais rápida. Em um ambiente de produção fabril, um lote que é rapidamente encontrado com algum defeito pode ser descartado mais rapidamente. As seguintes funções dos LIMS ajudam a aumentar a eficiência e eficácia:

1. Cálculos automáticos
2. Geração automática de relatórios
3. Validação dos dados de acordo com algum critério
4. Entrada automática de dados - o LIMS é capaz de acessar dados diretamente de instrumentos e/ou importar dados de arquivos gerados por instrumentos da base de dados do LIMS
5. Consulta

2.2 Boas Práticas de Laboratório

De acordo com o Inmetro [2012], *Good Laboratory Practices (GLP)* ou Boas Práticas de Laboratório (BPL) são o conjunto de orientações que tratam da organização, processos e condições sob as quais estudos de laboratório são planejados, executados, monitorados, registrados e relatados. A adoção das BPL tem como intenção promover a qualidade e validação dos resultados de pesquisa. As BPL são um sistema de qualidade aplicado a laboratórios que desenvolvem estudos, pesquisas e novas formulações, que necessitam da concessão de registros para comercialização, renovação ou

modificação de registros para produtos químicos, agrotóxicos, farmacêuticos, veterinários, cosméticos, alimentícios e monitoramento do meio ambiente e da saúde humana. As BPL têm como finalidade avaliar o potencial de riscos e toxicidade de produtos, objetivando a proteção da saúde humana, animal e do meio ambiente em vários casos, entre eles: em estudos que fundamentam a concessão, renovação ou modificações de registros de produtos, obtenção de propriedades físicas, químicas, físico-químicas e dados de segurança, petição para estabelecimento, modificação ou isenção de tolerância, estudos conduzidos em resposta a questionamentos de órgãos governamentais, estudos de campo, entre outros. Os laboratórios que desenvolvem pesquisa devem conduzir seus estudos segundo os princípios de BPL, de acordo com os critérios estabelecidos pela *OECD (Organization for Economic Cooperation and Development)*, pois resultados com qualidade comparáveis são a base de aceitação mútua entre países. Nesse sentido, os princípios de BPL são formalmente recomendados para serem usados pelos países membros da OECD desde 1981 [Embrapa, 2012].

2.3 Fluxo de Trabalho

Segundo a *Workflow Management Coalition* [Coalition, 2012] fluxo de trabalho corresponde a um processo de negócio onde documentos, informações ou tarefas são passadas de um participante para o outro para execução de uma ação global. As atividades de um fluxo de trabalho podem ocorrer concorrentemente e eventualmente impactar umas nas outras, de acordo com um conjunto de regras.

2.3.1 XML Process Definition Language

O *XML Process Definition Language (XPDL)* é um padrão definido pela *Workflow Management Coalition (WfMC)* para transferência de fluxo de trabalho entre produtos ou ferramentas localizados separadamente. Essa linguagem permite a separação entre o ambiente de desenvolvimento do fluxo de trabalho e o ambiente de execução do mesmo. Isto é, a definição de um processo gerado por uma ferramenta, pode ser usada como entrada por diferentes ferramentas/produtos.

O *XPDL* permite a representação de um fluxo de trabalho que seja capaz de manipulação automática, tal como modelagem por um sistema de gerenciamento de fluxo de trabalho. O processo de definição consiste em um encadeamento de atividades e seus relacionamentos, um critério que indique o início e término do processo e informações das atividades como participantes e dados [Coalition, 2012].

Essa linguagem de especificação usa o XML como mecanismo de especificação de fluxo de trabalho. O *XPDL* compreende um padrão de intercâmbio comum que permite aos produtos suportarem representações internas arbitrárias de definição de processos. Isso acontece graças à variedade de mecanismos usados para transferir dados entre sistemas de acordo com as características do cenário de negócio. Em todos os casos, a definição do fluxo de trabalho deve ser expressa em uma forma consistente, o que é possível em função de um conjunto comum de objetos, relacionamento e atributos expressando seu conceito. Os principais elementos para definição de fluxo de trabalho são os processos e suas atividades que possuem características definidas como atributos e transições (Figura 2.1).

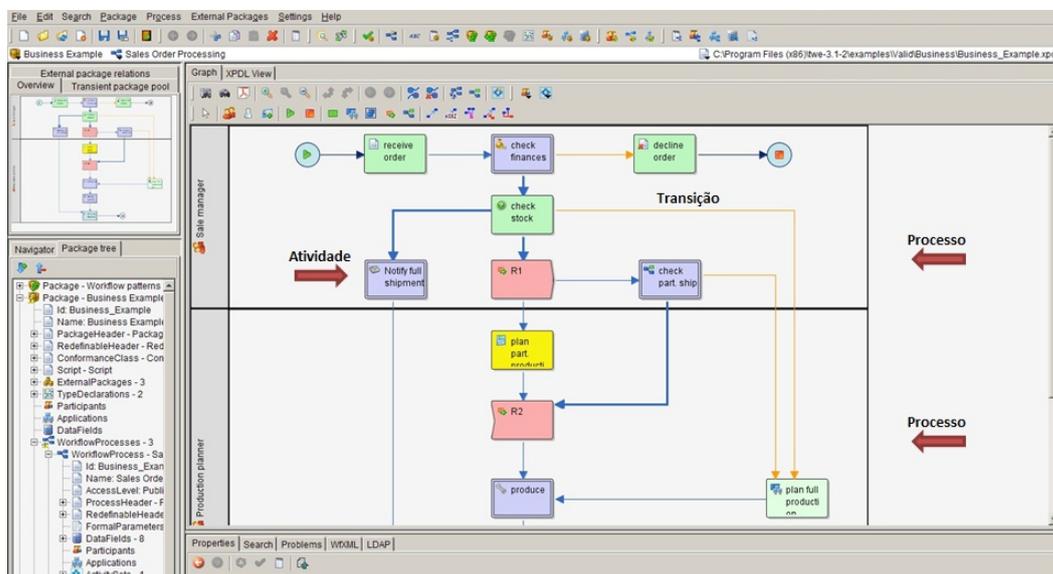


Figura 2.1. Elementos básicos de um fluxo de trabalho, processos, atividades e transição.

2.3.1.1 Processo

O processo no XPDL representa o processo no mundo real e contém as informações associadas com a administração como data de criação, autor, etc. Além disso, pode conter informações que podem ser usadas durante a execução do processo (parâmetros de inicialização, notificações, etc.). O processo é constituído por atividades.

2.3.1.2 Atividade

Um processo compreende uma ou mais atividades, cada uma compreendendo uma lógica e auto contida unidade de trabalho do processo. Uma atividade representa um

trabalho o qual será processado pela combinação de recursos especificados por um usuário ou uma aplicação. O escopo de uma atividade é local para um processo (fluxo de trabalho) específico.

Uma atividade pode ser especificada como uma repetição, ou seja, a atividade atua como controle de uma repetição da execução de parte do fluxo de trabalho dentro do processo.

2.3.1.3 Transição

Atividades são relacionadas entre si através de condições para controle do fluxo de trabalho. Estas condições são chamadas de transições. Cada transição tem três propriedades elementares: atividade-origem, atividade-destino e a condição com a qual a transição será realizada. A transição de uma atividade para a outra pode ser condicional ou incondicional. As transições do fluxo de trabalho podem resultar em uma operação sequencial ou paralela de atividades. As transições definem a forma como o fluxo de trabalho sofrerá ramificação ou convergência. Essa abordagem permite ao fluxo de trabalho controlar processamentos de atividades paralelas e assim ramificações e sincronizações podem ser gerenciados tendo as atividades associadas. O escopo de uma transição é local, ou seja, tem efeito somente dentro do fluxo de trabalho e as atividades associadas.

2.3.1.4 Aplicação

Aplicações externas ou interfaces podem ser invocadas pelo fluxo de trabalho para apoiar ou automatizar o processamento associado com cada atividade. As aplicações podem ser ferramentas industriais ou serviços corporativos. A definição de aplicações no fluxo de trabalho reflete a interface entre o fluxo de trabalho e a aplicação, incluindo quaisquer parâmetros requeridos.

2.3.1.5 Variáveis

Em um processo é possível definir variáveis que serão utilizadas durante sua execução. Para criar uma variável deve-se definir um identificador, um nome e se a variável é do tipo vetor ou não. Existem outros parâmetros que são opcionais. Por exemplo, para uma variável é possível atribuir-lhe uma descrição, definir o seu valor inicial padrão ou determinar o tamanho da variável. Além destes parâmetros é preciso também definir o tipo da variável. Os principais tipos que uma variável pode assumir são: booleano, data, alfanumérico, inteiro e real. Existem outras possibilidades de tipos como o tipo enumerado ou o tipo declarado.

2.3.1.6 Atributos estendidos

Algumas atividades requerem atributos estendidos para expressar características adicionais. Os atributos estendidos são definidos pelo usuário para expressar qualquer entidade de características adicionais. Os atributos estendidos podem ser adicionados a qualquer um dos componentes do XPDL (pacotes, processos, atividades, variáveis, etc.). Um atributo estendido é composto por um nome e um valor. Apesar de a especificação XPDL conter a maioria das construções que possam ser necessárias na definição de processos, pode haver circunstâncias em que informações adicionais do usuário deverão ser incluídas. Deste modo, quando as extensões são necessárias são usados os atributos estendidos.

2.3.1.7 Parâmetros Formais

Além de variáveis, um processo pode conter parâmetros formais, que são atributos passados durante a invocação do processo. Um parâmetro formal irá possuir um identificador, uma descrição e um modo que irá definir se trata de um parâmetro de entrada ou de saída. Como na definição de uma variável, para um parâmetro formal também deve ser definido um tipo.

2.3.2 Together XPDL Workflow Editor

Together Workflow Editor (TWE) [Editor, 2010] é um editor gráfico para a construção, edição e gerenciamento de arquivos de definição de processos baseados no WfMC XPDL. Essa ferramenta é baseada na versão 2.1 do formato xml XPDL publicado pela WfMC. O editor TWE simplifica o processo de criação e edição de arquivos XPDL, uma vez que representa todos os elementos do XPDL graficamente através de painéis e componentes gráficos, usando a notação gráfica BPMN. Isso permite que o usuário tenha melhor entendimento e visão geral da definição de processos. No TWE existem várias funcionalidades que permitem encontrar atividades específicas, participantes, aplicações, erros no modelo e outros. A saída final do editor é um arquivo xml (usando o padrão XDPL), o qual pode ser interpretado e executado por qualquer ferramenta que trabalha com o WfMC XPDL.

O TWE possui, portanto, três finalidades principais: ler arquivos no formato XPDL independentemente de qual ferramenta o arquivo se originou; representação gráfica e guia para edição/modelagem dos processos e escrita de arquivos xml com a definição de processos no formato XPDL [Teamsolutions, 2012].

2.4 Obtenção de Plantas Geneticamente Modificadas

O laboratório do Núcleo de Biologia Aplicada da EMBRAPA Milho e Sorgo - Sete Lagoas - Brasil - trabalha na obtenção de sementes de milho e sorgo geneticamente modificadas, ou seja, na obtenção de organismos transgênicos. Organismos transgênicos são aqueles nos quais o DNA é modificado usando engenharia genética com o objetivo de introduzir uma característica que não ocorre naturalmente naquela espécie. O laboratório da EMBRAPA utiliza duas técnicas para obtenção desses organismos: transformação via *Agrobacterium tumefaciens* e transformação via Biobalística.

2.4.1 Transformação via *Agrobacterium tumefaciens*

Este método de transformação utiliza um sistema natural de transferência de genes presente na bactéria *Agrobacterium*. *Agrobacterium* é uma bactéria de solo capaz de causar tumores vegetais na região da infecção. Estes tumores resultam da presença do plasmídeo Ti ou plasmídeo indutor de tumor na célula bacteriana. O plasmídeo Ti é uma molécula circular grande (200 a 800 kb) de DNA de fita dupla que pode se replicar independentemente do genoma de *Agrobacterium tumefaciens* [Gelvin, 2003]. Localizado no plasmídeo Ti, encontram-se duas regiões importantes para a transferência de genes da bactéria para a planta, a região do T-DNA e a região vir. As regiões dos T-DNAs de plasmídeos selvagens contêm genes que comandam a produção de opinas e hormônios, tais como auxina e citocinina, pela célula vegetal. As opinas são aminoácidos utilizados apenas por *Agrobacterium* como fonte de carbono e nitrogênio, enquanto os hormônios são responsáveis pela indução dos tumores vegetais. O T-DNA tem aproximadamente, entre 10 e 30 Kb e suas extremidades são delimitadas por duas sequências de 25 pb altamente homólogas, denominadas extremidades direita e esquerda. A *Agrobacterium* selvagem transfere o seu T-DNA através das membranas das células vegetais e o incorpora no DNA genômico da planta. O processamento do T-DNA e sua transferência para a célula vegetal é devido em grande parte a atividade de virulência das proteínas codificadas na região vir [Gelvin, 2003].

Agrobacterium tumefaciens constitui um excelente sistema de introdução de genes em células vegetais uma vez que: (i) o DNA pode ser introduzido em todos os tecidos da planta, o que elimina a necessidade da produção de protoplastos; (ii) a integração do T-DNA é um processo relativamente preciso. A região do DNA a ser transferida está definida pelas sequências flanqueadoras, extremidades direita e esquerda. Ocasionalmente se produzem reordenações, mas na maioria das vezes a região é inserida intacta

no genoma da planta. Normalmente, os T-DNA integrados mostram mapas genéticos consistentes e segregação adequados. Ademais, os caracteres introduzidos por esta via têm se mostrado estáveis durante muitas gerações de cruzamentos. Esta estabilidade é crítica quando se pretende comercializar as plantas transgênicas geradas [Yukoh et al., 1994; Yuji et al., 1996].

2.4.2 Transformação via Biobalística

O bombardeamento de células vegetais com DNA de interesse é um método direto de transformação desenvolvido para introduzir ácidos nucléicos no genoma ou plastoma de células [Nigel & Claude, 2002]. É uma metodologia comumente utilizada por laboratórios que trabalham com transformação genética de plantas. Foi desenvolvida no final dos anos 80 para manipular o genoma de plantas recalcitrantes à transformação mediada por *Agrobacterium*, dentre as quais são incluídos os cereais [Klein et al., 1987; Nigel & Claude, 2002].

Na transformação via bombardeamento de partículas ou biobalística, micropartículas de metal cobertas com o gene de interesse (GDI) são aceleradas em direção as células alvo, utilizando equipamentos conhecidos como “gen gun” ou canhão gênico [Klein et al., 1987; Sanford, 1988], com velocidades suficientes para penetrar a parede celular e não causar a morte da célula. DNA precipitado sobre as micropartículas é liberado gradualmente dentro da célula pós-bombardeamento, e integrado ao genoma [Klein et al., 1987; Nigel & Claude, 2002]. A aceleração das micropartículas é obtida por uma faísca de alta tensão elétrica, ou uma descarga de hélio. As partículas utilizadas devem ser não-tóxicas, não-reativas e menores que o diâmetro da célula-alvo. Normalmente são utilizadas micropartículas de ouro ou tungstênio. No método original, um sistema de aceleração a base de pólvora foi usada para impulsionar DNA revestido por partículas de tungstênio, que penetraram a membrana plasmática resultando na expressão gênica [Klein et al., 1987]. Dispositivos modernos como, por exemplo, PDS 1000 (BioRad Laboratories, Hercules, CA, E.U.A.) utilizam o gás hélio, ou no caso do gene gun Accell (Agracetus, Inc., Middleton, WI, E.U.A.) a eletricidade, para impulsionar micropartículas em direção a células alvo. Vários parâmetros físicos correlacionados com o equipamento de biobalística, tais como pressão, distância de voo do macrocarreador e dos microcarreadores, e vácuo, precisam ser otimizados para o sucesso da transformação. Além destes parâmetros, os biológicos relacionados ao material vegetal e ao GDI que será utilizado também devem ser estudados em experimentos preliminares [Sanford et al., 1993].

A partir da década de 90 a biobalística foi utilizada para transformar uma grande

variedade de plantas, inclusive o milho. Gordon-Kamm [1990] e Fromm et al. [1990] foram os primeiros grupos a relatarem a produção de milho transgênico a partir do bombardeamento de calos embriogênicos.

2.5 Trabalhos Relacionados

Atualmente é importante que pesquisadores, além de gastar horas em pesquisa, apresentem numerosas publicações e terem desenvolvido novas metodologias. É necessário demonstrar que todo o processo de análise experimental ocorreu com qualidade e de forma correta [Gimeno, 2003]. Para isto, os Sistemas de Gestão de Qualidade se fazem cada vez mais presentes em empresas e laboratórios. Apesar de existirem sistemas específicos para cada ramo de atividade, como a TS16949 (para indústrias de autopeças), a ISO9001 (para empresas em geral), as Boas Práticas de Laboratório (BPL) ou ISO/IEC17025 (para laboratórios), a estrutura básica de todos eles é muito semelhante, [Olivares, 2006]. Gimeno [2003] conseguiu resumir em 3 pontos os princípios da gestão de qualidade: 1) Documente o que faz, 2) Faça o que documenta e 3) Procure a qualidade do produto e do serviço que prometeu a seu cliente.

O emprego dos princípios da gestão de qualidade em laboratórios exige estratégias que garantam que todos os atores envolvidos sejam capazes de prezar por esses princípios. A literatura apresenta como solução o emprego dos Sistemas de Gerenciamento de Informação de Laboratórios (LIMS). Normalmente, os LIMS apresentados comercialmente são específicos de um ramo de atividade ou estão atrelados às especificidades de um laboratório. Sendo assim, existe o impedimento de que um mesmo LIMS seja usado por laboratórios diferentes. É comum que os laboratórios tenham procedimentos muito diferentes. Alguns exemplos de LIMS comerciais incluem o LabVantage Solutions [2010], Computing Solutions [2010] e o Solutions [2010]. Em Rasmussen et al. [2007] é apresentado um caso de estudo que consiste na implementação de um LIMS, UNICConnect. Esse é implementado com a preocupação de atuar sobre as muitas variáveis que podem comprometer a qualidade de um fluxo de trabalho em um laboratório. Num experimento para captura de informações celulares, o LIMS apresentado em Rasmussen et al. [2007] permite controle de algumas restrições sobre esse processo, como por exemplo o bloqueio do uso de reagentes com data de validade expirada e também o impedimento do uso de instrumentos fora do período de tempo de certificação. Além do mais, um e-mail de alerta é enviado para o técnico responsável pelo experimento. Outro trabalho demonstrando a utilização de LIMS pode ser visto em Balaji et al. [2006], onde há a implementação de um LIMS com aplicação na genotipagem. Em um

alto nível de abstração o projeto deste LIMS pode ser dividido em quatro principais módulos: configuração do experimento - módulo que permite ao usuário definir o nome do projeto, experimento e a entrada dos dados necessários ao experimento; acompanhamento da amostra - esse módulo permite o acompanhamento de todo o fluxo de trabalho da genotipagem de uma amostra de DNA; geração de relatórios - dados submetidos a base de dados poderão ser apresentados ao usuários na forma de relatórios; e armazenamento de dados - todas as informações a cerca da extração de DNA, e os protocolos são armazenados nesse módulo.

Apesar do significativo número de LIMS disponíveis nos dias atuais, seu uso ainda é limitado em alguns laboratórios já que cada laboratório tem necessidades específicas. Laboratórios de produção e análise de transgênicos estão entre os laboratórios que não encontram inúmeros LIMS disponíveis no mercado capaz de atender as peculiaridades dos processos que envolvem a produção de organismos transgênicos. Um exemplo de sistema deste tipo é o *Villager (Villager Transgenic Animal Management System)*, uma solução proprietária para o gerenciamento da informação relacionada com o desenvolvimento de animais geneticamente modificados [Inc., 2011]. Atualmente não há sistema similar para laboratórios que trabalham com plantas transgênicas e a maioria das empresas e laboratórios trabalhando nessa área desenvolvem seus sistemas exclusivos que não são considerados LIMS ou apresentam severas limitações. Kohl & Gremmels [2010], em um recente trabalho, discutem a importância do registro organizado dos dados dos experimentos de transformação de plantas e propõe um sistema para documentação desse tipo de serviço e pesquisa mas essa ferramenta não é um verdadeiro LIMS.

Capítulo 3

Visões

3.1 Visão no Contexto de Banco de Dados

Em Sistemas Gerenciadores de Banco de Dados, o conceito de visão é uma tabela derivada de outras tabelas sendo que essas outras tabelas podem ser tabelas da base ou visões (*views*) previamente definidas. Uma visão não necessariamente existe em forma física, ela é considerada uma tabela virtual ao contrário da tabela da base [Elmasri & Navathe, 2011]. As visões podem prover vantagens sobre as tabelas da base, isto é, as visões podem representar um subconjunto de dados contidos numa tabela; podem esconder a complexidade dos dados e podem limitar o grau de exposição dos dados.

Uma visão pode ser considerada como um modo de especificar uma tabela que precisa ser referenciada com frequência mesmo não existindo fisicamente. As visões em banco de dados são muito solicitadas, por exemplo, quando se quer limitar o acesso aos dados de uma tabela por determinados usuários. Há cenários em que não é permitido que um usuário X tenha acesso a todas as colunas de uma tabela T , dessa forma uma visão é uma maneira interessante de estabelecer uma tabela virtual com apenas os dados (colunas) que esse usuário pode enxergar ou manipular.

3.2 Visão no Contexto de LIMS

Nesse trabalho, o conceito de visão é baseado no encontrado em banco de dados. Visões, nesse trabalho, são formas diferentes de enxergar os dados que não seguem o fluxo de trabalho experimental. E também podem ser agrupamentos de alguns dados experimentais (subconjunto destes dados) que têm alguma propriedade em comum que não necessariamente é diretamente relacionada ao fluxo de trabalho.

Esse trabalho propõe que fluxos de trabalhos sejam desenvolvidos de forma que mesmo fluxos de trabalho com uma relação de dependência, um fluxo de trabalho X é responsável pelo processo de preparo de um insumo que é utilizado pelo fluxo de trabalho Y , possam ser executados em momentos diferentes. Dessa forma tem-se uma visão, ou seja, há um fluxo de trabalho para execução de processos experimentais e outro fluxo de trabalho que faz parte do processo, porém seu fluxo de dados não segue o fluxo principal do processo experimental. A Figura 3.1 ilustra isso, dado que o fluxo de trabalho A da figura é um fluxo de trabalho de experimento, nesse caso, transformação via *Agrobacterium tumefaciens*, os fluxos B e C são visões dos dados do fluxo A . O fluxo B é o fluxo de trabalho de agendamento, ou seja, esse fluxo de trabalho representa uma agenda que o usuário define estabelecendo quantos dias após a execução de uma atividade, a atividade posterior deverá ser executada. O fluxo B é uma visão do fluxo A (transformação via *A. tumefaciens*). O fluxo A usa os dados estabelecidos no fluxo B . Assim como o fluxo B o fluxo C também é uma visão, porém uma visão que se refere ao preparo de solução. As soluções preparadas por esse fluxo serão usadas no fluxo experimental.

Enquanto tem-se a visão como uma forma diferente de visualizar aqueles dados que não seguem o fluxo principal de dados experimental, tem-se também a visão relativa a um subconjunto de dados do fluxo de dados experimental. Esse subconjunto tem características em comum não atreladas ao fluxo de trabalho como um todo e que o discrimina do fluxo principal de trabalho experimental. Na Figura 3.1, no fluxo de trabalho A , os triângulos coloridos que correspondem às atividades A_4 , A_6 , A_8 e A_{10} podem ser analisadas (vistas) separadamente do restante do fluxo de trabalho. Esses triângulos representam **Entradas e Saídas**, são um conjunto de dados que mesmo analisados em separado trazem significado. Em um exemplo, as entradas e saídas que vão da atividade *Transferência para o Primeiro Meio de Seleção* até a atividade de *Destinação*. Se o usuário olhar apenas essas entradas e saídas ele é capaz de entender a evolução de um embrião até se tornar uma semente transgênica, vendo a sequência das atividades. Ou seja, o embrião tornou-se um calo, depois tornou-se uma planta, depois a planta foi polinizada, colhida e por fim a semente foi estocada.

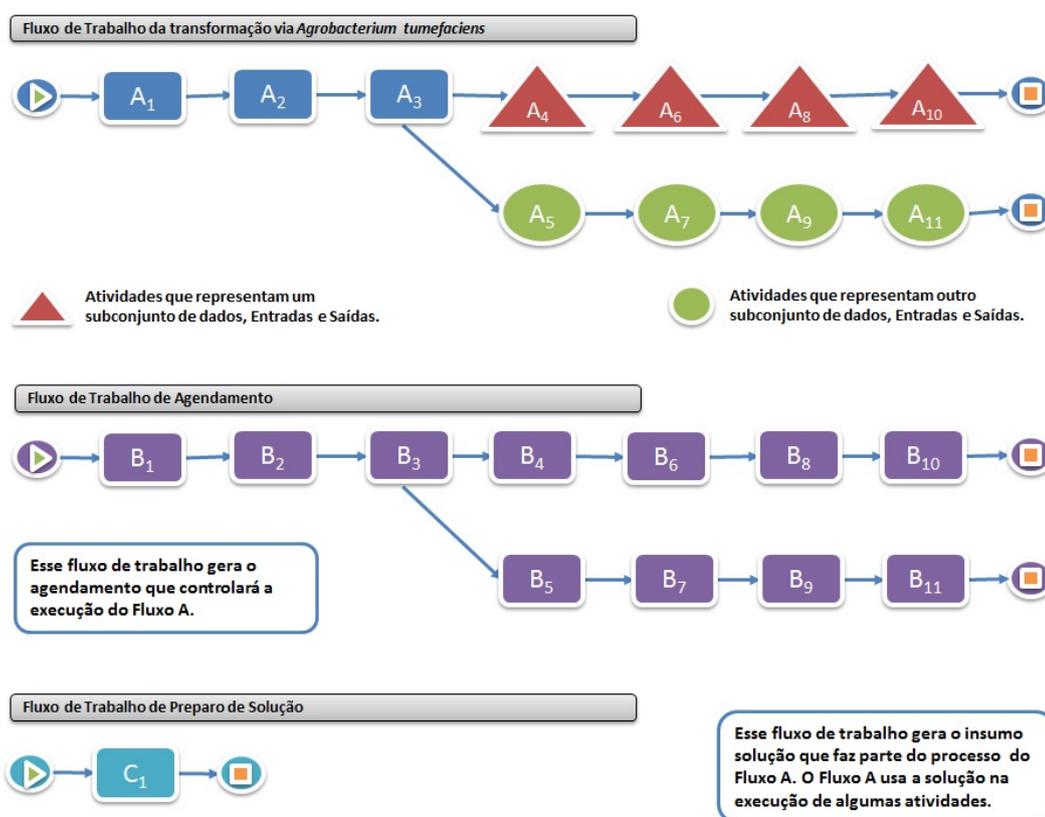


Figura 3.1. Exemplo de fluxo de trabalho com visões. O fluxo de trabalho A é o fluxo experimental enquanto os fluxos B e C são as visões de agendamento e controle de insumos, respectivamente. No fluxo A, os triângulos e os círculos coloridos em destaque representam subconjuntos de dados considerados visões do fluxo.

3.3 Visão de Controle de Insumos

Nesse trabalho, a entidade insumo é extremamente importante para o entendimento do controle de insumos, haja vista que essa entidade é o objeto central do controle de insumos. Entende-se como insumo, uma entidade física que é usada pelos laboratórios para realização de algum experimento, por exemplo, reagentes para preparo de meio de cultura para realização dos experimentos de transformação por *A. tumefaciens* ou por *Biobalística*. Um insumo tem uma característica extremamente importante, sua quantidade. A quantidade é o que diferencia a entidade insumo de outra entidade qualquer como, por exemplo, a entidade paciente no contexto de um consultório médico.

Em um laboratório é comum o manuseio de insumos e há referência aos mesmos sob duas perspectivas, a de uma informação genérica e a de uma outra informação específica. Em geral, quando o pesquisador compra reagentes ou construções gênicas ele recebe uma caixa com os mesmos. Quando prepara soluções de meios de culturas

ou outras, estas podem ser aliquotadas em frascos. Inicialmente, os reagentes são todos iguais, existindo apenas informações globais dos mesmos. Entretanto, no dia-a-dia do laboratório, de acordo com o uso dos reagentes, os mesmos serão discriminados sendo associado a eles dados particulares.

Ao se referir a um insumo, faz-se referência ao objeto ou conjunto de objetos físicos que existem no laboratório. Sendo assim, entende-se como informação macro, a informação genérica do conjunto de objetos. A informação específica, por sua vez, se refere aos dados de cada objeto do conjunto. A figura 3.2 ilustra a diferença entre a informação macro e a específica. Na figura, a caixa representa a informação genérica que é compartilhada pelos frascos 1, 2, 3, 4, 5 e 6. Exemplos de informação macro seriam o nome do insumo, o lote e as datas de fabricação e validade. A informação específica é o que discrimina os frascos, apresentados na figura, uns dos outros. Ela terá a numeração do frasco e pode acontecer, por exemplo, da quantidade de cada frasco variar de um para o outro à medida que o reagente é usado para realizar um experimento. A quantidade é alterada e essa informação pertence apenas ao frasco usado e não aos demais. Adicionalmente, se um reagente acabar ou o frasco for danificado, o mesmo pode ser declarado como descartado. Neste caso, apenas o frasco em questão terá essa informação e não os outros.

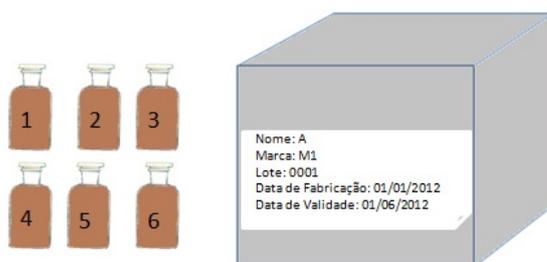


Figura 3.2. Exemplo de informação macro e específica relativas a um insumo. A informação macro é representada pela caixa e a informação específica por cada um dos frascos da mesma (1, 2, 3, 4, 5 e 6).

A visão de controle de insumos permite a definição de quatro tipos de fluxos de trabalho: cadastro, preparo e visualização de um insumo e cadastro de receita. Esses fluxos de trabalhos juntos formam a visão de controle de insumos. A Figura 3.3 apresenta os componentes da visão assim como a relação entre eles. Inicialmente, tem-se o cadastro de um insumo (por exemplo, insumo *A*, *B*, *C* e *D*). Os insumos cadastrados serão armazenados no banco de dados *MySQL* (seta 1). Após o cadastro, é possível fazer o preparo de um novo insumo, ou seja, os insumos existentes, já cadastrados,

poderão ser usados para o preparo deste novo insumo (seta 2). Os insumos existentes no banco de dados podem ser acessados pelo fluxo de trabalho *Preparo de insumo* (seta 3).

O componente *receita* está diretamente relacionado com o elemento *Preparo de insumo* (seta 4). Para preparar um insumo, é necessário uma receita, essa receita tem o nome de todos os insumos necessários para fazer outro insumo. A receita guiará o usuário no preparo, uma vez que diminuirá possíveis erros de entrada de dados (escolha de insumos). Como a receita tem a lista de insumos que deverão ser usados no preparo, o usuário sempre ficará alerta sobre quais usar. Após fazer o preparo de um insumo, esse irá também para o banco de dados (seta 5).

Os fluxos de trabalho do controle de insumos são ortogonais aos fluxos de trabalho de realização de experimentos. Ser ortogonal nesse caso indica que esses fluxos de trabalho tem um ponto de interseção apenas e que os fluxos são independentes. O ponto de interseção é o uso dos insumos pelos fluxos de trabalho para realização de algum experimento, ou seja, no momento da execução de um experimento, os insumos serão acessados pelo mesmo, haja vista que estão no banco de dados (seta 6).

Além do cadastro e preparo de insumos, o controle de insumos tem um fluxo de trabalho que permite a visualização dos insumos. Isso permite ao usuário ter acesso a situação do estoque, isto é, a quantidade em estoque, os insumos que estão em falta, etc. A grande importância desse componente é permitir o planejamento do laboratório quanto a reposição e manutenção do estoque. Descartar insumos fora da validade de uso e viabilizar o consumo daqueles próximos da data de validade.

3.3.1 Cadastro

Um dos componentes da visão de controle de insumos é o fluxo de trabalho **cadastro**. A razão para a existência deste fluxo são as limitações atuais existentes no cadastro de insumos feito pelos pesquisadores sem uma ferramenta automática. É comum no dia-a-dia de um laboratório a atividade de catalogar os insumos usados para realizar um experimento qualquer, como por exemplo, transformação via *A. tumefaciens*. Geralmente o pesquisador compra várias caixas de variados tipos de insumos e em seguida cataloga esses de alguma forma, seja em registros em papel ou em planilhas eletrônicas. O cadastro dos insumos em arquivos de papel torna inviável a geração de relatórios e análises de dados. A utilização de planilhas eletrônicas tem limitações, uma vez que estas não foram projetadas como um LIMS ou parte integrante de um. Além disso, os insumos a serem cadastrados podem variar em seus atributos de acordo com o tipo de experimento a ser executado. A utilização do fluxo de trabalho cadastro responde a es-

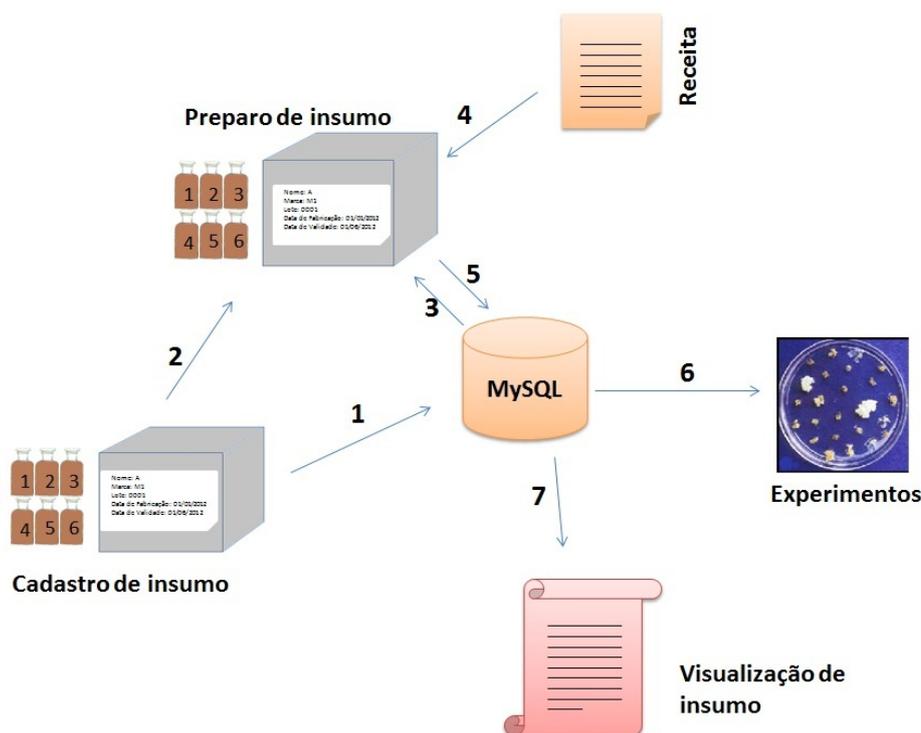


Figura 3.3. Elementos que compõem a visão de controle de insumos. As setas indicam a existência de um fluxo de informação de um componente para o outro. 1. Armazenamento do insumos cadastrados no banco de dados (BD); 2. Preparo de novos insumos com outros já cadastrados; 3. Acesso aos insumos no BD pelo fluxo de trabalho *Preparo de Insumos*; 4. Relacionamento entre os componentes *Receita* e *Preparo de Insumo*; 5. Inserção de novo insumo no BD; 6. Acesso dos insumos do banco de dados na execução de experimentos; 7. Visualização dos insumos que foram cadastrados ou preparados e registrados no banco de dados.

tas questões, uma vez que automatiza o processo de cadastro e como o fluxo é definido na ferramenta TWE, isto torna a definição do cadastro de insumos flexível o suficiente para ser adotado por vários laboratórios para cadastro de insumos de experimentos diferentes.

Na construção do fluxo de trabalho **Cadastro de Reagentes** (Figura 3.4), duas informações são mandatórias, o nome da tabela onde ficará armazenada a informação macro e o nome da tabela onde será armazenada a informação específica. Com essas duas informações, o *framework* Flux, ao ler a definição do fluxo de trabalho, é capaz de gerar os formulários para o cadastro. Outras configurações podem ser feitas na definição do fluxo de trabalho. Para isso existe um conjunto de características que podem ser atribuídas aos campos do formulário de cadastro. Um exemplo disto é aquele quando um campo é preenchido usando opções de um *drop down* menu, dependendo do valor escolhido outros campos podem ou não aparecer. Esse comportamento e outros

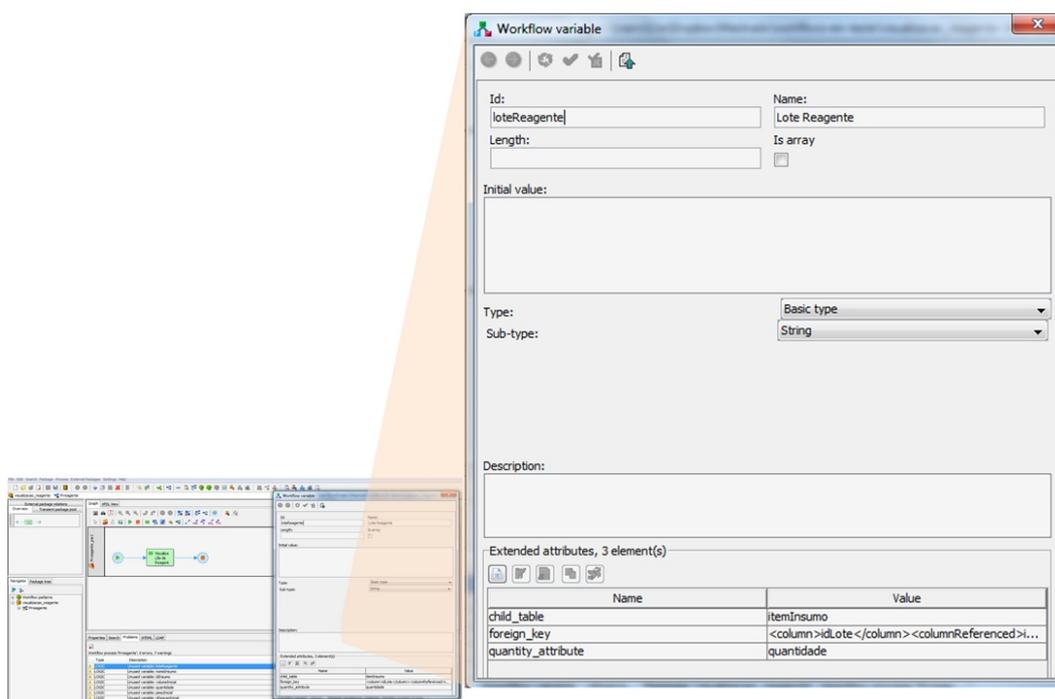


Figura 3.4. Construção do fluxo de trabalho **Cadastro de Reagente** no Together Workflow Editor.

similares são associados ao cadastro na definição do fluxo de trabalho.

3.3.2 Cadastro de Receita

Para a realização de um procedimento em laboratório existe um conjunto de etapas e informações essenciais que constituem um protocolo de laboratório. Estes protocolos possuem a descrição detalhada dos procedimentos e insumos necessários para a realização do experimento. Assim, faz parte da visão controle de insumos a definição do **Cadastro de receita**. Esse padrão para definição de fluxos de trabalho permite o cadastro de um protocolo. O protocolo pode ser qualquer informação na forma de campos de formulário que represente os procedimentos, por exemplo, nome do protocolo, arquivo (pdf, doc, etc) com a descrição e qualquer outra informação relevante. No mesmo **Cadastro de Receita**, o usuário deve colocar a lista de insumos que serão usados para preparar outro insumo (Figura 3.5). Assim, a receita cadastrada nesse componente do controle de insumos será usada como entrada para o preparo de insumos.

Este cadastro, assim como os outros tem os dois blocos de informações macro e específicos. Como exemplo, a Figura 3.5 apresenta a dinâmica desse cadastro no qual o usuário preenche o nome e anexa o protocolo da receita. Além disso, acrescenta a lista dos insumos necessários à receita. Por fim, o resultado do uso dessa é apresentado no



Figura 3.5. Exemplo do *Cadastro de Receita*. O usuário insere um nome para identificar a receita, um arquivo em pdf que contem as instruções (protocolo) de preparo do *Acido Propionico* e a lista de insumos que serão necessários para preparar o insumo.

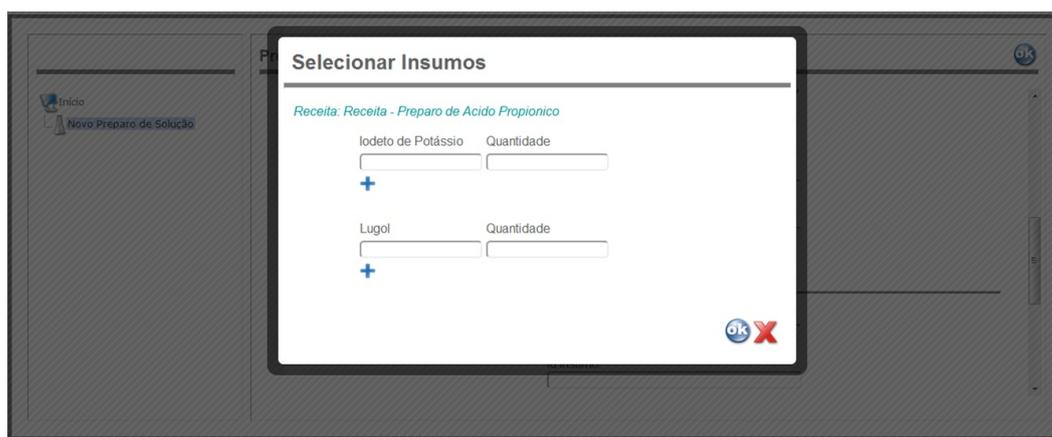


Figura 3.6. Resultado do uso da receita cadastrada na Figura 3.5.

exemplo da Figura 3.6. Essa tela é gerada quando o usuário vai preparar um insumo, a tela é apresentada de acordo com a criação da receita no **Cadastro de Receita**. Observa-se que a receita guia o usuário no momento de escolher os insumos e registra a quantidade que será usada e se o mesmo escolher uma quantidade inválida, o sistema faz as devidas verificações.

3.3.3 Preparo de Insumos

Geralmente o cadastro de insumos se dá nos casos em que os insumos são adquiridos do mercado através de uma compra ou de outros laboratórios através de empréstimo ou doação. Entretanto, existem casos em que o insumo é produzido no próprio laboratório,

ou seja, o pesquisador usa os insumos cadastrados, por exemplo, reagentes, para fazer um outro insumo. Esse processo de preparo de insumo carece de rastreamento para fins de auditoria e gerenciamento do que é consumido no laboratório, o que está em desacordo com a adoção das boas práticas de laboratório (BLP). Com o objetivo de tratar da automatização do registro deste processo e a adoção das BPLs, o controle de insumos contém o componente **Preparo de Insumos**.

O componente **Preparo de Insumos** apresenta os dois conjuntos de informações vistos na descrição do **Cadastro de Insumos** (informação macro e específica). A diferença entre estes componentes é que o **Preparo de Insumos** tem associado a si a receita para o preparo do insumo além do cadastro de informações presente no **Cadastro de Insumos**. Na definição do fluxo de trabalho **Preparo de Insumos**, deve ser especificado que este fluxo de trabalho terá um campo onde o usuário escolherá o protocolo a partir do qual o preparo será realizado (Figura 3.7). Com essa definição, na execução do fluxo de trabalho, o usuário terá uma tela (Figura 3.8) onde poderá escolher os insumos para o preparo.

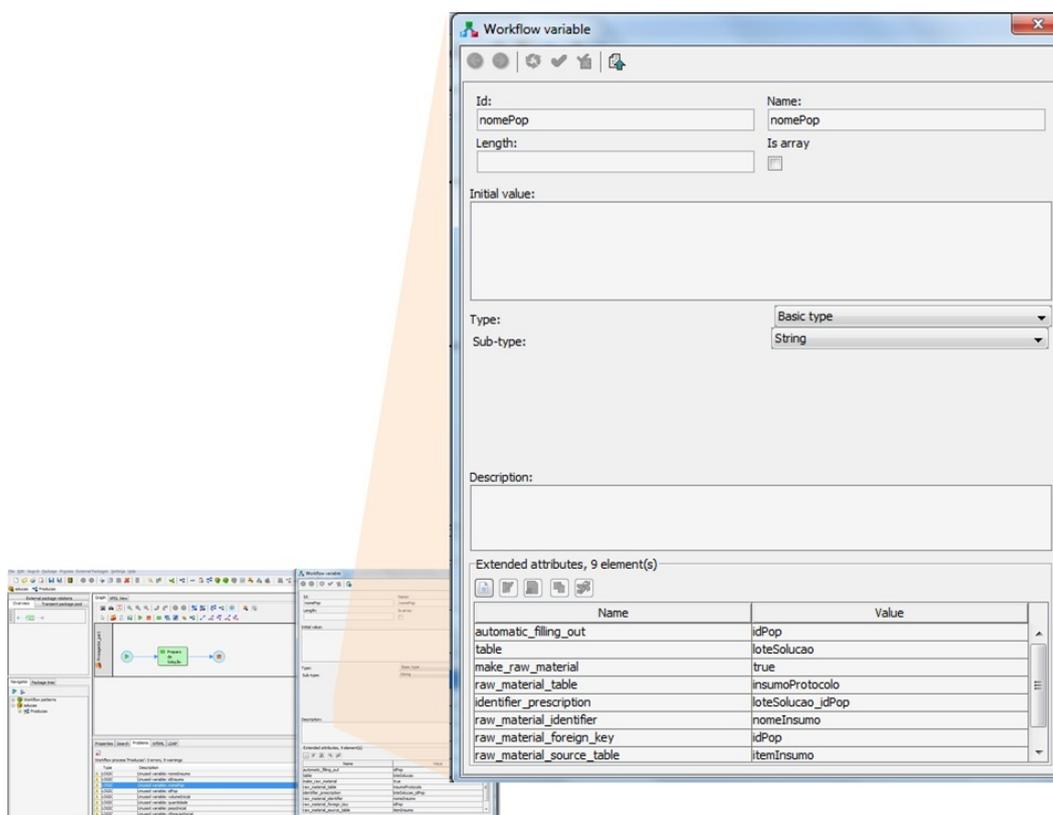


Figura 3.7. Construção do fluxo de trabalho **Preparo de Insumo** na ferramenta *Together Workflow Editor*. A definição do atributo *nomePop* se converterá no campo onde o usuário escolherá o protocolo para preparo da solução.

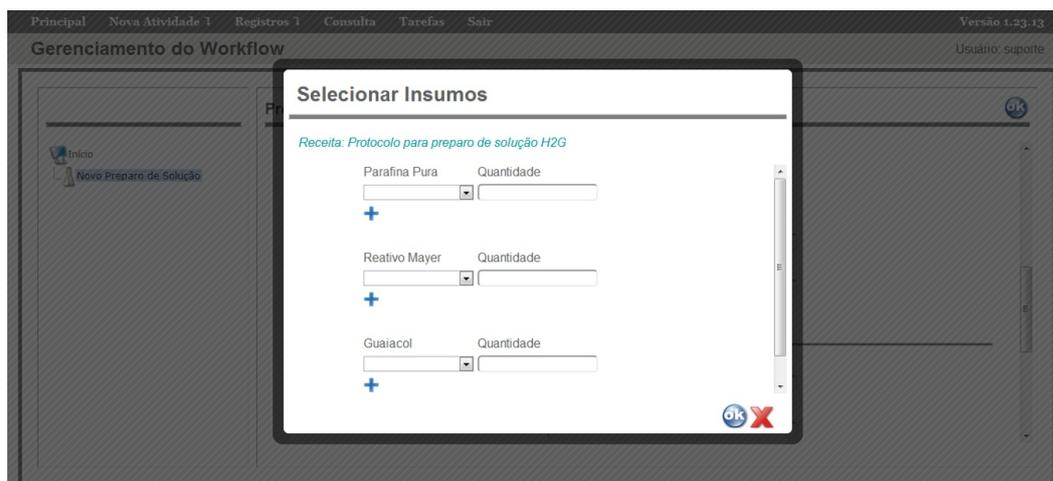


Figura 3.8. Tela da visão **Preparo de Insumos** onde o usuário escolhe os insumos a serem usados no preparo do insumo.

3.3.4 Visualização

A **Visualização** é outro componente da visão **Controle de Insumos**. Esse componente tem por objetivo apresentar a relação dos insumos catalogados no sistema de forma a permitir ao usuário acompanhar a dinâmica de entrada e saída dos insumos, ou seja, a quantidade atual global do estoque e a quantidade atual de cada tipo de insumo. Este componente permite ao gerente do laboratório planejar os insumos que deverão ser comprados e em que quantidade. Com este componente é possível, além disso, fazer pesquisas pelos insumos utilizando de filtros (Figura 3.9) e gerar relatórios da pesquisa em arquivos pdf e/ou planilha Excel.



Figura 3.9. Interface de visualização e edição dos insumos cadastrados e/ou preparados. Através do filtro o usuário pode fazer pesquisas variadas pelos insumos.

Neste fluxo de trabalho o usuário pode editar e excluir os insumos catalogados.

Porém, para excluir ou editar um insumo, o usuário deve ter um perfil especial de acesso ao sistema, ou seja, apenas usuários com privilégios de acesso podem fazer alterações no estoque de insumos. Assim, de acordo com os privilégios do usuário identificado no sistema, as opções editar e excluir estarão disponíveis ou não.

Esse componente é um fluxo de trabalho assim como os outros componentes (Cadastro, Preparo e Receita), portanto, a ferramenta TWE é utilizada para definir esse fluxo. Para esse fluxo de trabalho, informações como a informação macro (tabela com a informação macro) e a informação específica (tabela com a informação específica) são definidas no fluxo de trabalho.

3.3.5 Utilização dos Insumos

A existência da visão do controle de insumos permite que o laboratório tenha controle dos insumos que são usados nos experimentos executados no dia-a-dia do laboratório. Da mesma forma que é importante o controle da entrada de insumos no laboratório através dos cadastros e preparo, também é importante o controle do uso desses insumos. No **Preparo de Insumos** se tem o uso de insumos, isto é, ao preparar um insumo, utiliza-se um insumo previamente catalogado no sistema. Nos fluxos de trabalho experimentais, o uso de um insumo é feito da seguinte forma:

- Na definição do fluxo de trabalho no *Together Workflow Editor*, define-se qual insumo será usado, por exemplo, reagente;
- Ao executar uma atividade que exige uso de algum insumo, o usuário abrirá uma tela onde poderá escolher o insumo;
- O sistema apresenta para o usuário um campo onde o mesmo pode pesquisar pelo insumo em questão. O insumo será identificado por um nome e um identificador;
- Além de escolher o insumo, o usuário entra com a quantidade que será consumida;
- O sistema valida a quantidade inserida pelo usuário e associa o insumo e a quantidade escolhida pelo usuário a atividade em questão;

3.3.6 Conclusão

A visão controle de insumos atende os principais elementos envolvidos no gerenciamento dos insumos utilizados no dia-a-dia de um laboratório de pesquisas experimentais. Permite a entrada dos insumos através do módulo de cadastro e preparo de insumos, além disso, assegura através da receita que os procedimentos para preparo de um insumo

sejam devidamente seguidos pelo pesquisador ou outra pessoa no momento do preparo. Isso impede possíveis erros de entrada de dados em um momento de distração do usuário. A visão em questão assegura também que os insumos possam ser consumidos por outros processos (fluxo de trabalho) fazendo a correta calibração do estoque, o que permite ao gerente do laboratório conhecimento da dinâmica do que é consumido no laboratório. Além do mais, o módulo de visualização da visão controle de insumos permite que o responsável pelo laboratório tenha acesso a relatórios sobre o estado do laboratório em termos do que existe no estoque do mesmo e possibilita também a atualização do estoque sendo que essa atualização está sujeita ao nível de privilégios que o usuário possui.

3.4 Visão Entradas e Saídas

No fluxo de trabalho definido no TWE e interpretado pelo sistema Flux, as informações inerentes às atividades são armazenadas na forma de atributos. Desta forma, uma atividade tem um conjunto de atributos que assegura significado à atividade. Além dos atributos, outras informações associadas a atividade são aquelas denominadas *Entradas* e *Saídas*. Uma atividade terá uma *saída*, quando essa atividade — que faz parte de um experimento — produz um ativo que poderá ser usado ou consumido por uma atividade seguinte, ou seja utilizado como uma *entrada*. No caso dos experimentos do Laboratório da EMBRAPA, tem-se, por exemplo, como ativos, calos que são gerados em uma atividade e serão usados na seguinte. A necessidade de construir uma visão de entradas e saídas é que na versão anterior a esse trabalho, as entradas e saídas existiam, porém, era apenas possível trabalhar com entradas e saídas atômicas. Isso quer dizer que ao se cadastrar uma saída, essa será usada como entrada na atividade seguinte, por inteiro. Isso não corresponde à realidade dos laboratórios, incluindo o laboratório da EMBRAPA. Para suprir essa necessidade, foi construída a visão **Entradas e Saídas**, com a capacidade de lidar com entradas e saídas atômicas e não atômicas.

3.4.1 Saída Atômica

Uma saída atômica é aquela que será usada por uma próxima atividade por inteiro. Isto é, se em um experimento, uma atividade gera como saída um tubo com algum ativo, todo o ativo desse tubo deverá ser usado na próxima atividade. Essa característica de uma saída se enquadra nos cenários onde a saída sempre será usada como entrada na próxima atividade, sem que essa seja particionada de alguma forma. Nesse caso, não é possível ter um cenário no qual uma saída seja usada por duas atividades seguintes. A

Figura 3.10 mostra um exemplo de saída atômica em que existem quatro atividades, sendo que, por exemplo, uma saída da atividade A (S_{A1}) poderá se tornar entrada apenas para uma das atividades seguintes, B ou C . No exemplo da figura S_{A1} será entrada da atividade B .

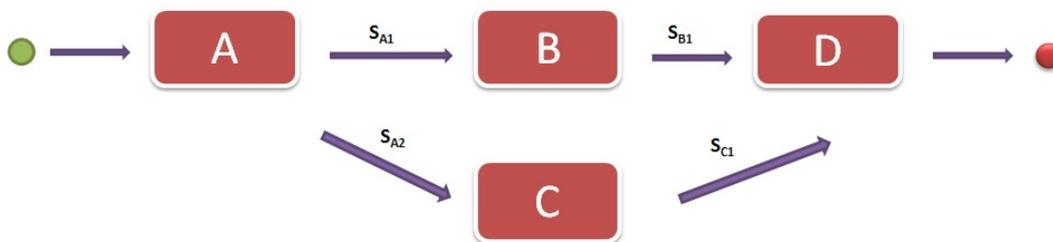


Figura 3.10. Exemplo de saída atômica na visão **Entradas e Saídas**.

3.4.2 Saída Não-atômica

Ao contrário da saída atômica, uma saída não-atômica é aquela que será apresentada para a próxima atividade de forma que seja possível determinar a quantidade a ser usada. Por exemplo, ao executar a atividade *Transferência Para Segundo Meio de Seleção* (Figura 3.11) o usuário têm várias opções de entradas com o respectivo campo para preencher a quantidade e a visualização da quantidade disponível. Sendo assim, uma saída gerada por uma atividade anterior pode ser usada por mais de uma atividade subsequente. Este comportamento se aplica em especial aos experimentos com transgênicos, haja vista que a atividade *Transferência Para Primeiro Meio de Seleção* pode gerar embriões como saídas e esses serão usados como entrada tanto para a atividade *Transferência Para Segundo Meio de Seleção* quanto para a atividade *Transferência Para Meio de Maturação*.

O exemplo da Figura 3.12 ilustra os aspectos da saída não-atômica. Veja que a atividade A tem a saída S_{A1} e que essa saída será entrada tanto para a atividade B quanto para a C . Isso é possível porque B receberá parte da saída, ou seja, cinco unidades da mesma e C também receberá parte da saída, isto é, cinco unidades.

3.4.3 Rastreamento das Saídas

Além da visão **Entradas e Saídas** permitir que se associem às atividades entradas e saídas atômicas e não atômicas, esta permite também que se relacione uma saída com uma entrada. Quando uma saída é cadastrada, em alguns casos essa saída é resultado da alteração do estado de alguma entrada. Pode-se citar como exemplo o cenário no

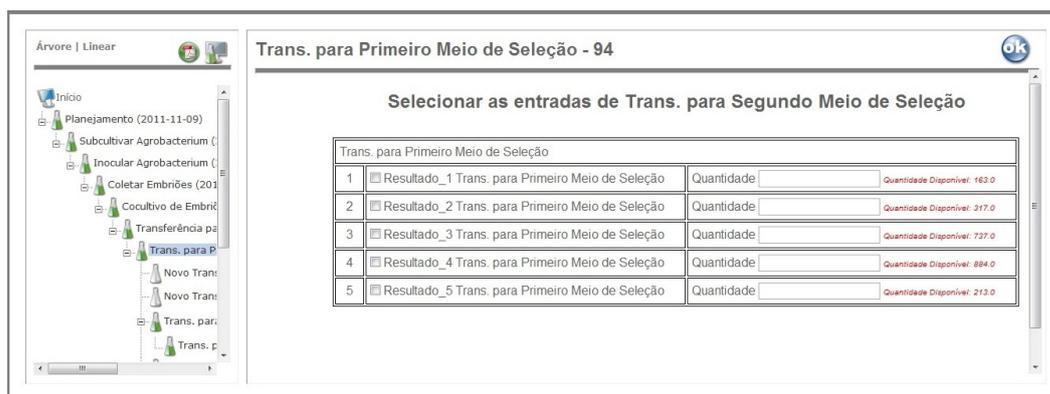


Figura 3.11. Interface da visão **Entradas e Saídas** para seleção das entradas a serem usadas na próxima atividade.

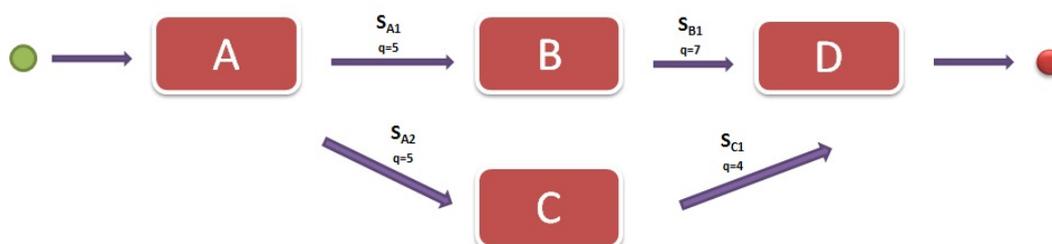


Figura 3.12. Exemplo de saída não atômica na visão **Entradas e Saídas**.

qual a atividade *Transferência Para Segundo Meio de Seleção* (Figura 3.13) gera um calo $C2$ como saída. Este calo é resultado da entrada do calo $C1$ que veio da atividade *Transferência Para Primeiro Meio de Seleção*, ou seja, o calo $C1$ se dividiu gerando mais um calo ($C2$). Assim, a saída $C2$ está associada com a entrada $C1$. A informação dessa relação é importante para fins de auditoria e rastreamento, inclusive, no contexto de aquisição de plantas transgênicas.

3.4.4 Conclusão

Com a visão entradas e saídas, os fluxos de trabalho se tornam mais robustos uma vez que passam a permitir um novo tipo de entrada/saída, não-atômica. Isso aproxima o LIMS Flux ainda mais da realidade dos laboratórios. Além do mais, o rastreamento das saídas das atividades dos fluxos de trabalho se apresenta como uma ferramenta poderosa a serviço de auditorias sendo que isso é comum quando se quer provar que os procedimentos e ativos produzidos em laboratórios são de qualidade.

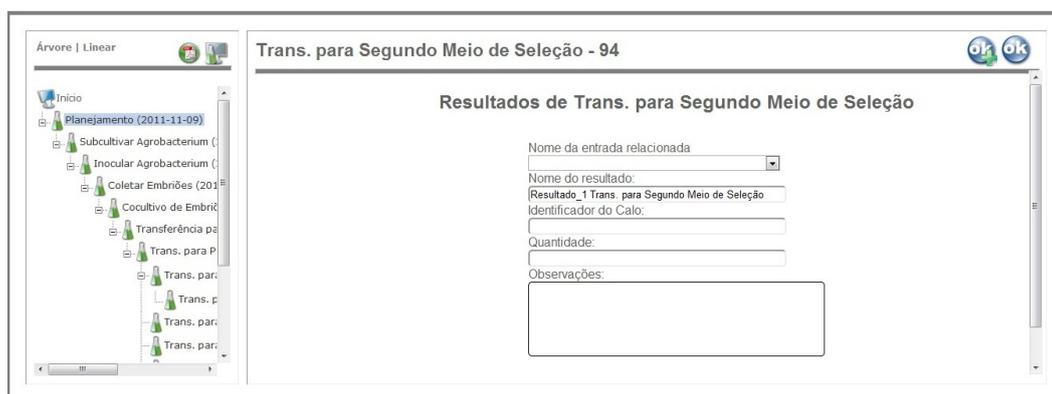


Figura 3.13. Interface da visão **Entradas e Saídas** para escolha das entradas a serem usadas na atividade.

3.5 Visão de Agendamento

No cotidiano de um laboratório, espera-se que vários fluxos de trabalhos sejam executados paralelamente. Em alguns laboratórios, todos os dias, fluxos de trabalhos são iniciados. Isso implica que ter-se-ão inúmeras atividades no sistema disponíveis para execução, algumas com intervalos de tempo de execução maiores outras menores. Diante disso é importante que o LIMS auxilie o pesquisador a lidar com fluxos de trabalho com dinâmicas de execução completamente diferentes. Portanto, a Visão **Agendamento** tem por objetivo permitir mecanismos de alerta e escalonamento de execução de atividades no tempo.

A visão **Agendamento** é um fluxo de trabalho que é derivado automaticamente do fluxo principal de execução de experimentos. O fluxo de agendamento é praticamente uma cópia do fluxo principal (Figura 3.10), porém, enquanto no fluxo principal cada atividade tem um conjunto de atributos que estão estritamente relacionados com a função da atividade, no fluxo de agendamento todas as atividades tem dois atributos: *Quantidade de dias* e *Ação para execução fora do prazo*.

O atributo *Quantidade de dias* se refere a quantidade de dias após a execução da atividade anterior que a atividade deverá ser executada. Veja o fluxo de trabalho *X Agendamento* (Figura 3.10). Caso atribua ao atributo *Quantidade de dias* da atividade *B* o valor 10, significa que a atividade *B* do fluxo *X* será executada dez dias após a execução da atividade *A* desse mesmo fluxo. Diante disso, o sistema listará na tela de tarefas após dez dias que a atividade *B* está disponível para execução.

No fluxo de trabalho de agendamento se define qual o comportamento das atividades do fluxo principal quando essas forem executadas fora da data estimada de execução. Sendo assim, o atributo *Ação para execução fora do prazo* apresenta uma

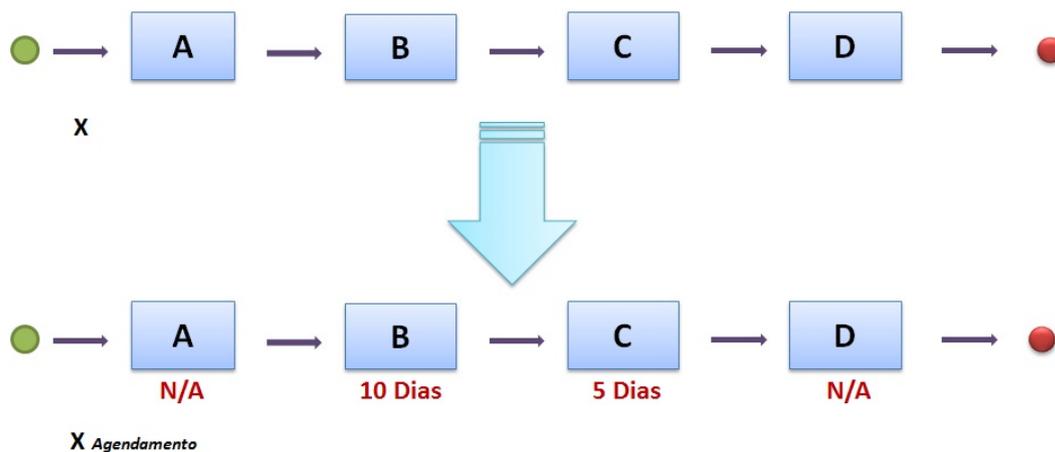


Figura 3.14. O fluxo de agendamento (X Agendamento) é gerado a partir do fluxo principal (X).

lista *drop down* com as opções de comportamento:

- *Bloquear* - Essa opção determina que a atividade do fluxo principal será bloqueada caso a execução não seja no prazo determinado, ou seja, o usuário não será capaz de executar a atividade e nem continuar com o fluxo de trabalho.
- *Alertar* - Essa opção implica em um alerta para o usuário quando o mesmo tenta executar a atividade do fluxo principal antes ou após a data estipulada de acordo com a informação submetida no atributo *Quantidade de dias*.
- *N/A* - Essa opção não provoca nenhuma ação quando a atividade do fluxo principal é executada fora do prazo.

O agendamento permite a organização dos pesquisadores quanto ao escalonamento das atividades a serem executadas. É comum que um mesmo pesquisador inicie a execução de vários fluxos de trabalho ao mesmo tempo. Logo, o agendamento é fundamental para alertar e liberar a execução de atividades na data correta evitando assim que a execução de determinados passos (atividades) de um experimento não seja negligenciado. Assim o responsável por um experimento não corre o risco de executar uma atividade fora do prazo comprometendo o sucesso do experimento. Além do mais, como o agendamento de um fluxo de trabalho é comum para todas as instâncias do fluxo de trabalho, então a visão Agendamento permite a calibração do agendamento (escalonamento) das atividades, isto é, a primeira instância do fluxo de trabalho executada acompanhará o escalonamento definido. Porém, caso esse escalonamento não seja o ideal, o pesquisador é capaz de verificar isso durante a execução do fluxo e alterar

o agendamento de forma que as próximas execuções (instâncias) do fluxo de trabalho sejam contempladas com um agendamento mais próximo do ideal.

Capítulo 4

Implementação

As visões foram implementadas no sistema Flux sendo que esse LIMS é uma aplicação Web e dessa forma o desenvolvimento das visões fez uso da tecnologia *Java (J2EE)* de tal forma que o componente *Servlet* que faz parte dessa tecnologia é o responsável pelo processamento de requisições e respostas *HTTP*. O *Servlet* é um componente que atua como servidor responsável em gerar *HTML* e *XML* para a camada de apresentação de um aplicativo Web. Enquanto o componente *Servlet* é o responsável pelas requisições e respostas dos clientes, o componente *Java ServerPages* [Oracle, 2012] permite a criação de páginas web dinâmicas. Esse componente é o responsável pela interface das aplicações Web, isso é possível pelo fato da sintaxe do *JSP* ser uma mistura de dois conteúdos básicos: *scriptlet* e *markup*. *Markup* é tipicamente um padrão HTML ou XML, enquanto os *scriptlet* são blocos de código Java os quais podem ser unidos com o *markup*. Esses elementos juntos é o que permite a dinamicidade das páginas web baseadas nessa tecnologia.

O MySQL [MySQL, 2012] é o Sistema de Gerenciamento de Banco de Dados do LIMS FLUX. A modelagem do banco é baseada principalmente em dois elementos: **Tabelas Básicas** e **Tabelas Externas**. As **Tabelas Básicas** se referem às tabelas criadas no decorrer da construção da aplicação e que são necessárias independentemente do laboratório que o LIMS irá gerenciar, ou seja, essas tabelas são o núcleo do modelo de dados do LIMS. A tabela *workflow* (Figura 4.1) é um exemplo de tabela básica. Nessa tabela são registrados os fluxos de trabalhos construídos na ferramenta *Together Workflow Editor* e que foram lidos pelo LIMS para gerar os formulários (atividades) e suas relações para gerenciar processos de um laboratório. O conjunto de **Tabelas Básicas** corresponde a um total de 56 tabelas. Diferentemente, as **Tabelas Externas** estão estritamente relacionadas com o laboratório e seus processos que a aplicação irá gerenciar. As **Tabelas Externas** junto com os fluxos de trabalhos são responsáveis

pela flexibilidade do LIMS. Essas tabelas são definidas no momento da modelagem do laboratório.

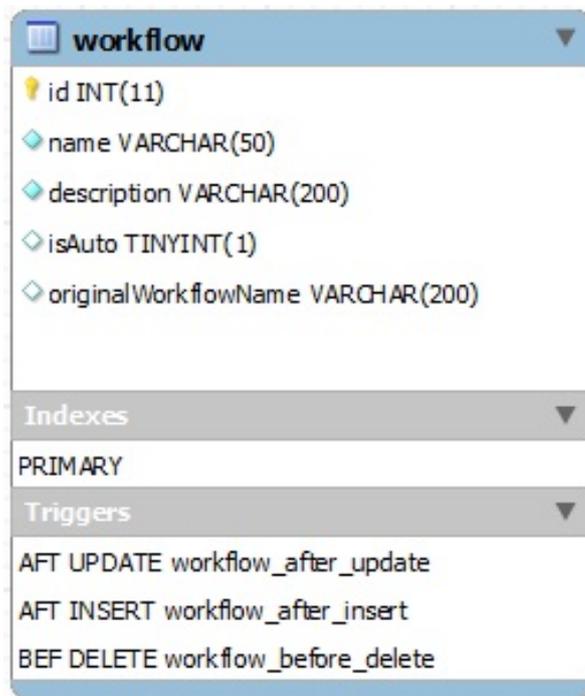


Figura 4.1. Tabela básica *workflow*, nessa tabela é gravado informações a cerca dos fluxos de trabalhos gerenciados pelo Flux.

O Tomcat é o servidor Web do LIMS Flux, Este é um servidor Web Java, mais especificamente, um container de *servlets* [Foundation, 2012]. O Tomcat é um servidor de aplicações *JEE* desenvolvido pela *Apache Software Foundation*. É distribuído como software livre, sendo oficialmente adotado pela Sun como a implementação de referência para as tecnologias *Java Servlet* e *JavaServer Pages* (JSP). Esse servidor inclui ferramentas para configuração e gerenciamento, o que também pode ser feito editando-se manualmente arquivos de configuração formatados em *XML*.

4.1 Visão de Controle de Insumos

Os elementos fundamentais da implementação do controle de insumos são: arquivo no formato *XPDL* com o fluxo de trabalho e o modelo ER implementado no banco de dados. No fluxo de trabalho estão definidas as tabelas da informação macro e específica sendo que essas tabelas fazem parte do modelo ER de **Tabelas Externas**. De posse desses elementos, o LIMS Flux, através de um parse sobre o arquivo *XPDL*, identifica as tabelas onde serão mapeadas as informações macro e específica. Com o nome das

tabelas, o sistema tem acesso ao meta dados das mesmas, ou seja, as propriedades de cada coluna da tabela (nome, tipo, etc). Dessa forma, o Flux tem controle das informações necessárias para gerar os formulários, seja de cadastro ou preparo de um insumo. O XPDL especifica a relação de um para muitos entre a tabela da informação macro e a tabela da informação específica. Dessa forma, o sistema Flux apresenta os formulários em conformidade com esse aspecto, ou seja, ao fazer um cadastro ou preparo de insumos, quando o sistema insere essas informações no banco de dados, a informação específica passa a ter uma chave estrangeira da informação macro.

Os campos que aparecem nos formulários (fluxo de trabalho) do controle de insumos correspondem às colunas das tabelas macro e específica. Porém, como o nome das colunas das tabelas do banco de dados geralmente não é apropriado para ser o rótulo do campo do formulário, então o sistema Flux disponibiliza um módulo onde o usuário pode associar um rótulo a cada campo de uma tabela externa específica. Essa configuração fica armazenada nas tabelas básicas *recordTable* e *recordAttribute* (Figuras 4.2 e 4.3) e ao gerar os formulários o sistema consulta essas tabelas para obter os rótulos correspondentes a cada coluna das tabelas.

4.2 Entradas e Saídas

A implementação das entradas e saídas considera a existência do arquivo *XPDL* com a especificação do fluxo de trabalho e nesse fluxo de trabalho as definições que se referem às entradas e saídas. O algoritmo que lida com essa visão espera dois tipos de entradas e saídas: saída atômica e saída não-atômica. Além disso, saída pode ser acoplada com o mecanismo de rastreamento, ou seja, é possível registrar a entrada que originou a saída. Quando a saída é atômica, a mesma é registrada no banco e quando essa saída se tornar uma entrada para uma outra atividade, ela será associada pela atividade na sua totalidade. Porém, quando uma saída se tratar de uma saída não-atômica, isso implica que a saída será registrada obrigatoriamente com um campo que corresponderá a uma quantidade. O sistema ao verificar que a saída é não-atômica, apresenta o campo que corresponderá a quantidade e o valor desse campo será gravado na coluna *quantity* da tabela básica *result* (Figura 4.4). Quando a saída não-atômica se torna uma entrada para determinada atividade, a entrada é apresentada para o usuário de tal forma que o mesmo entre com o valor da quantidade da entrada que a atividade utilizará. Dessa forma, caso o usuário não solicite toda a entrada (valor inserido pelo usuário é menor que o valor da quantidade), o novo valor da entrada será maior que zero e sendo assim a entrada poderá ser usada por outra atividade. Isso explica o porquê da saída ser

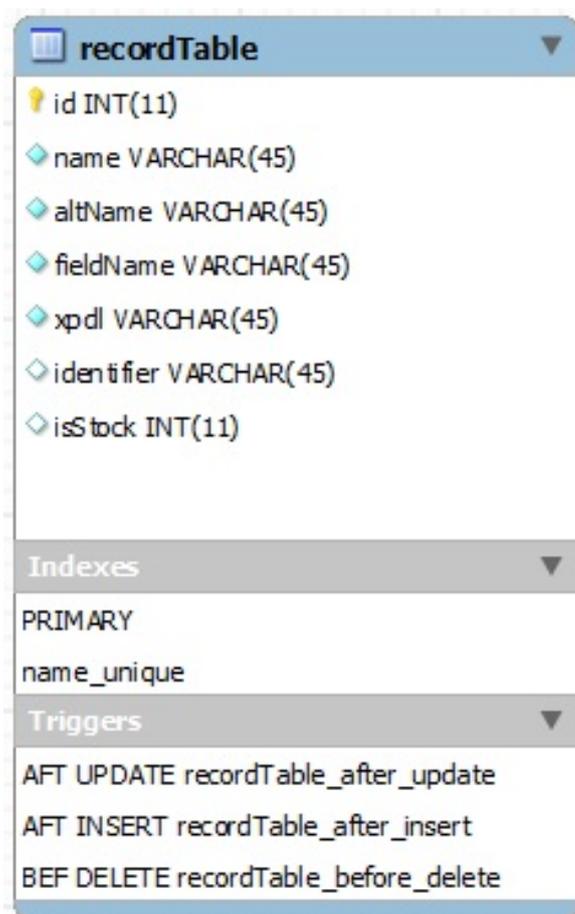


Figura 4.2. Nessa tabela básica são armazenados os meta-dados das tabelas externas.

denominada não-atômica.

A saída também pode ser rastreada, para isso o sistema apresenta no momento do cadastro da saída um *drop-down* com a lista das entradas da atividade para que o usuário indique qual é a entrada que originou ou está de alguma forma relacionada com a saída. As colunas *relatedInputId* e *relatedInputName* registram os dados da entrada relacionada com a saída. Com isso é possível acompanhar a evolução de um ativo pelo sistema.

4.3 Agendamento

A implementação da visão Agendamento requer pouca especificação de fluxo de trabalho. O agendamento é derivado de um fluxo de trabalho principal, geralmente o fluxo de trabalho responsável pelo experimento dentro do laboratório. O sistema Flux nota que um fluxo de trabalho tem agendamento quando está definido no fluxo de trabalho

The screenshot shows a database table named 'recordAttribute'. The columns are listed as follows:

- id INT(11)
- idRecordTable INT(11)
- name VARCHAR(45)
- type VARCHAR(45)
- notNull VARCHAR(6)
- minValue FLOAT
- maxValue FLOAT
- form at VARCHAR(45)
- altName VARCHAR(45)
- xpdl VARCHAR(45)
- idRecord INT(11)
- displayTotal TINYINT(4)

The table also has the following indexes:

- PRIMARY
- fk_attr_record
- fk_attribute_record

The table also has the following triggers:

- AFT UPDATE recordAttribute_after_update
- AFT INSERT recordAttribute_after_insert
- BEF DELETE recordAttribute_before_delete

Figura 4.3. Tabela básica onde são gravados os meta-dados das colunas das tabelas externas.

o atributo *SCHEDULE* (Figura 4.5). Dessa forma, o seguinte algoritmo é executado:

1. Verificar se o fluxo de trabalho tem o atributo *SCHEDULE*;
2. Se tem, então é gerado uma cópia do fluxo de trabalho;
3. É gerado no banco de dados um novo registro na tabela *workflow* (Figura 4.1). Nesse novo registro a coluna *name* recebe o nome do fluxo de trabalho original concatenado com a cadeia de caracteres “_auto”. A coluna *isAuto* é preenchida com o numeral “1” e a coluna *originalWorkflowName* é preenchida com o nome do fluxo de trabalho original;

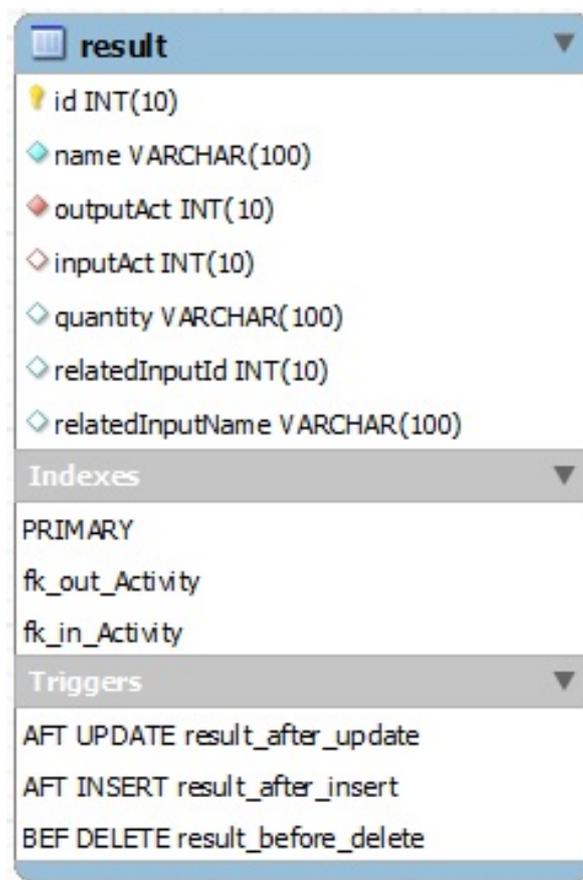


Figura 4.4. As informações das entradas e saídas são armazenadas nessa tabela básica.

Após o registro do novo fluxo de trabalho, o usuário terá na tela a lista de fluxos de trabalhos com o novo fluxo de trabalho gerado automaticamente para registrar o escalonamento das atividades do fluxo de trabalho original. Ao executar o fluxo de trabalho de agendamento os seguintes passos são executados pelo LIMS Flux:

1. Buscar as atividades do fluxo de trabalho a partir do nome do fluxo de trabalho original (coluna *originalWorkflowName*);
2. Identificar qual a atividade que será executada;
3. Se for a primeira atividade, associar os seguintes atributos à atividade: *Identificador*, *Quantidade de dias* e *Ação para execução fora do prazo*, caso contrário, associar todos os atributos citados com exceção do atributo *Identificador*;
4. Apresentar para o usuário o formulário da atividade a ser executada com os campos (atributos) associados;

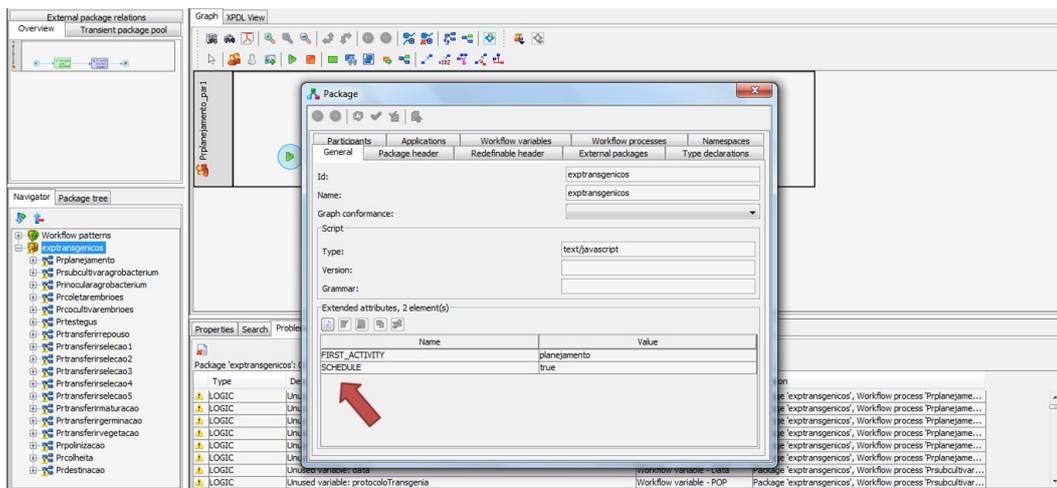
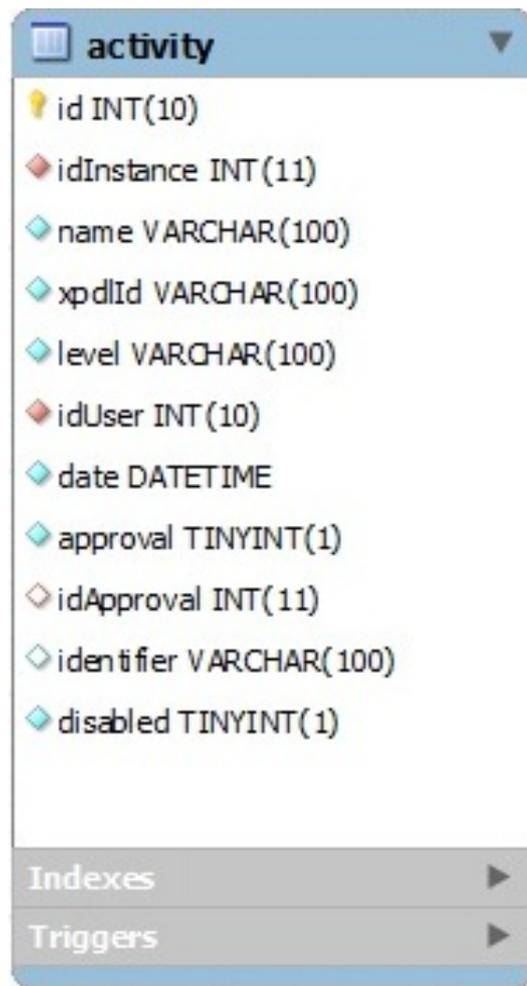


Figura 4.5. Especificação de um fluxo de trabalho na ferramenta *Together Workflow Editor* sendo que o mesmo possui agendamento.

Ao final do preenchimento dos campos do formulário e conclusão da execução pelo usuário, os atributos da atividade em vez de ser armazenados na tabela básica *activity* (Figura 4.6), são armazenados na tabela básica *schedule* (Figura 4.7).



The image shows a screenshot of a database management tool displaying the structure of a table named 'activity'. The table has the following columns and data types:

| Column Name | Data Type |
|-------------|--------------|
| id | INT(10) |
| idInstance | INT(11) |
| name | VARCHAR(100) |
| xpdId | VARCHAR(100) |
| level | VARCHAR(100) |
| idUser | INT(10) |
| date | DATETIME |
| approval | TINYINT(1) |
| idApproval | INT(11) |
| identifier | VARCHAR(100) |
| disabled | TINYINT(1) |

Below the column list, there are sections for 'Indexes' and 'Triggers', both with right-pointing arrows, indicating that these sections are currently collapsed.

Figura 4.6. Tabela onde são armazenados os dados das atividades.

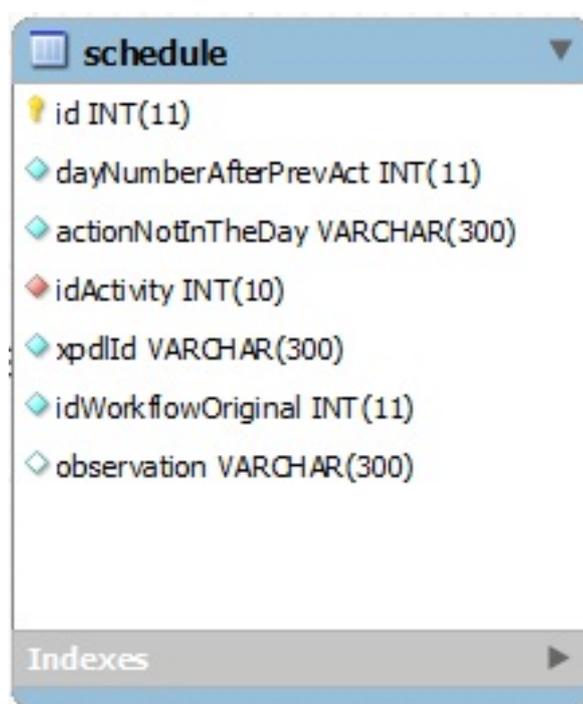


Figura 4.7. A tabela *schedule* faz parte do modelo de dados de tabelas básicas e armazena as informações referentes ao agendamento.

Capítulo 5

FluxTransgenics - Caso de Estudo

O *Laboratório de Biologia Celular e Cultura de Tecidos* da EMBRAPA Milho e Sorgo desenvolve trabalhos em biologia aplicada na identificação e desenvolvimento de marcadores moleculares para a seleção assistida (Transgenia) e para a identificação de genes de interesse econômico em milho e sorgo. Informações sobre a expressão e a função desses genes ampliam o conhecimento científico sobre os mecanismos moleculares e fisiológicos envolvidos na adaptabilidade e produtividade dessas culturas. Fazem parte do processo de transgenia duas técnicas de transformação, uma que utiliza a bactéria *Agrobacterium tumefaciens* e a outra a Biobalística. O primeiro método de transformação se baseia em um sistema natural de transferência de genes presente na bactéria *Agrobacterium* enquanto, no segundo, micropartículas de metal cobertas com o gene de interesse são atiradas em direção a célula que será modificada. Estas transformações são processos complexos realizados em laboratórios que devem ser executados de acordo com normas específicas, os princípios de BPL.

Quando se pensa na obtenção de uma planta transgênica como produto comercial. “Boas Práticas de Laboratório” ou BPL é um sistema da qualidade relativo ao processo organizacional e às condições sob as quais estudos não-clínicos referentes à saúde e meio ambiente são planejados, realizados, monitorados, registrados, arquivados e relatados. As BPLs são aplicadas a todos os estudos de segurança ambiental e de saúde exigidos pelos órgãos regulamentadores, visando o registro ou licença para produtos farmacêuticos, praguicidas, cosméticos, veterinários, aditivos de alimentos e rações e produtos químicos industriais, excetuando-se aqueles que sejam dispensados por legislação (Norma NIT-DICLA-035/INMETRO) [Estevam dos Santos, 2010]. Com o objetivo de se adequarem a essas práticas, os laboratórios do Núcleo de Biologia Aplicada (NBA) da EMBRAPA Milho e Sorgo [Embrapa, 2010] têm realizado diversas modificações em suas rotinas. Entre estas se inclui a necessidade de implantar um

sistema automático para gerenciamento de dados e experimentos realizados. Portanto, surgiu a necessidade de trabalhar com um LIMS para auxiliar no registro de suas atividades, controle de materiais e insumos, consultas e geração de relatórios. Embora LIMS sejam usados desde a década de 80 e tenham atendido a vários laboratórios clínicos, de construção e engenharia, pesticidas, resíduos e ambientais, ainda existem poucos desses sistemas como produto comercial disponível e os que existem são específicos para determinada aplicação, sendo ainda menor o número daqueles que contemplam as necessidades de um laboratório de produção de organismos transgênicos. Sendo assim, o Laboratório de Universalização de Acesso em conjunto com o Núcleo de Biologia Aplicada da EMBRAPA Milho e Sorgo iniciou um trabalho de construção de um LIMS para gerenciamento de todo o processo de criação de transgênicos da EMBRAPA Milho e Sorgo.

O Laboratório de Universalização de Acesso (LUAR) já possuía experiência anterior no desenvolvimento de LIMS, tendo desenvolvido o LIMS SIGLa [Simões et al., 2010], como resultado do trabalho de mestrado do aluno Alexandre Simões de Melo. O SIGLa é um LIMS baseado em fluxos de trabalho adaptáveis com suporte a múltiplos laboratórios. É uma aplicação web de fácil uso e alto grau de usabilidade que permite o gerenciamento de laboratórios com procedimentos completamente diferentes. A capacidade de adaptabilidade do SIGLa se deve em parte ao fato de que todos os procedimentos e atividades, ou seja, todo o fluxo de trabalho do laboratório é modelado externamente utilizando a ferramenta *Together XPDL Workflow Editor*, sendo gerado um arquivo XPDL com todas as definições necessárias para que o sistema, ao ler o arquivo, possa ser utilizado para gerenciar o laboratório.

Entretanto, o sistema SIGLa carece de algumas funcionalidades que são importantes para tornar a adaptabilidade mais eficiente. Assim, foi desenvolvido um novo sistema a partir deste para suprir estas necessidades, tendo o código fonte do SIGLa como ponto de partida. O sistema, chamado Flux, tem todas as funcionalidades do SIGLa, bem como o fato de ser baseado em fluxos de trabalho, mas inclui melhorias e extensões, inclusive, o componente de visões. O principal diferencial do FLux frente aos outros LIMS existentes é a capacidade de adaptar a vários tipos de laboratórios, haja vista que é baseado em fluxos de trabalho. Com a incorporação de visões no Flux, é possível se trabalhar com seis tipos de fluxos de trabalho. São eles: fluxo principal, cadastro de insumos, preparo de insumos, visualização de insumos, cadastro de receita, e agendamento. Como exemplo de uso prático desse LIMS, foi desenvolvido o LIMS *FluxTransgenics*. Esse sistema foi projetado para gerenciar dados e processos do laboratório de produção de transgênicos da EMBRAPA Milho e Sorgo - MG - Brasil. O *FluxTransgenics* é uma otimização do sistema Flux. É composto por dois

fluxos de trabalho para representar todos os passos da transformação de milho e sorgo (via *Agrobacterium tumefaciens* e Biobalística). Há também os fluxos de trabalhos de subcultivo de embriões e calos, plantio de sementes e subcultivo de controle positivo. Fazem parte do *FluxoTransgenics* também os fluxos de trabalhos da visão de controle de insumos e os fluxos da visão de agendamento. Para ilustrar a relação desses fluxos de trabalhos um com o outro, a Figura 5.1 mostra um exemplo onde se tem o fluxo de trabalho principal *Agrobacterium tumefaciens* com as visões relacionadas. O triângulo da figura representa a visão de Entradas e Saídas. Nota-se que essa visão atua apenas em um subconjunto de dados do fluxo de trabalho, ou seja, as atividades *Trans. para o 1º Meio de Solução* e *Trans. para a Casa de Vegetação*. Já as visões de Agendamento e Controle de Insumos atuam em todo o fluxo de trabalho. Por exemplo, o fluxo de trabalho de preparo de soluções da visão de controle de insumos gera insumos que são utilizados por várias atividades da transformação via *Agrobacterium tumefaciens*.

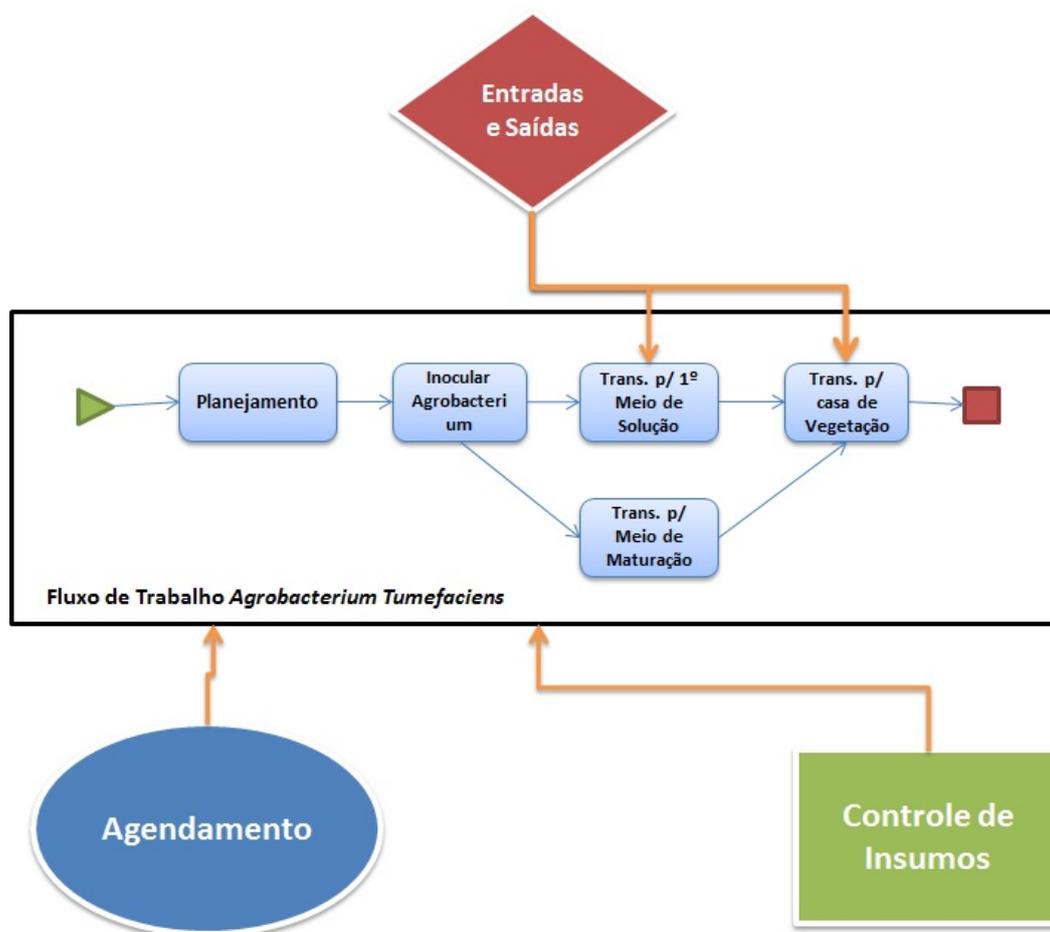


Figura 5.1. Abstração do relacionamento do fluxo de trabalho *Agrobacterium tumefaciens* com as visões Controle de Insumos, Agendamento e Entradas e Saídas.

5.1 Fluxo de Trabalho da Transformação via *Agrobacterium tumefaciens*

Esse fluxo de trabalho se refere à transformação via *Agrobacterium tumefaciens*. Este método de transformação utiliza um sistema natural de transferência de genes desenvolvido pela bactéria *A. tumefaciens*. *A. tumefaciens* é uma bactéria de solo capaz de causar tumores vegetais na região da infecção. Estes tumores resultam da presença do plasmídeo Ti ou plasmídeo indutor de tumor na célula bacteriana. O plasmídeo Ti é uma molécula circular grande (200 a 800 kb) de DNA fita dupla que pode se replicar independentemente do genoma de *A. tumefaciens*. Localizado no plasmídeo Ti se encontram duas regiões importantes para a transferência de genes da bactéria para a planta, a região do T-DNA e a região vir. O plasmídeo é usado para inserir genes na planta que não são nativos desta. O fluxo de trabalho que representa o experimento de transformação por *A. tumefaciens* considera todos os passos da transformação, a qual tem vinte passos representados por vinte atividades (Figura 5.2).

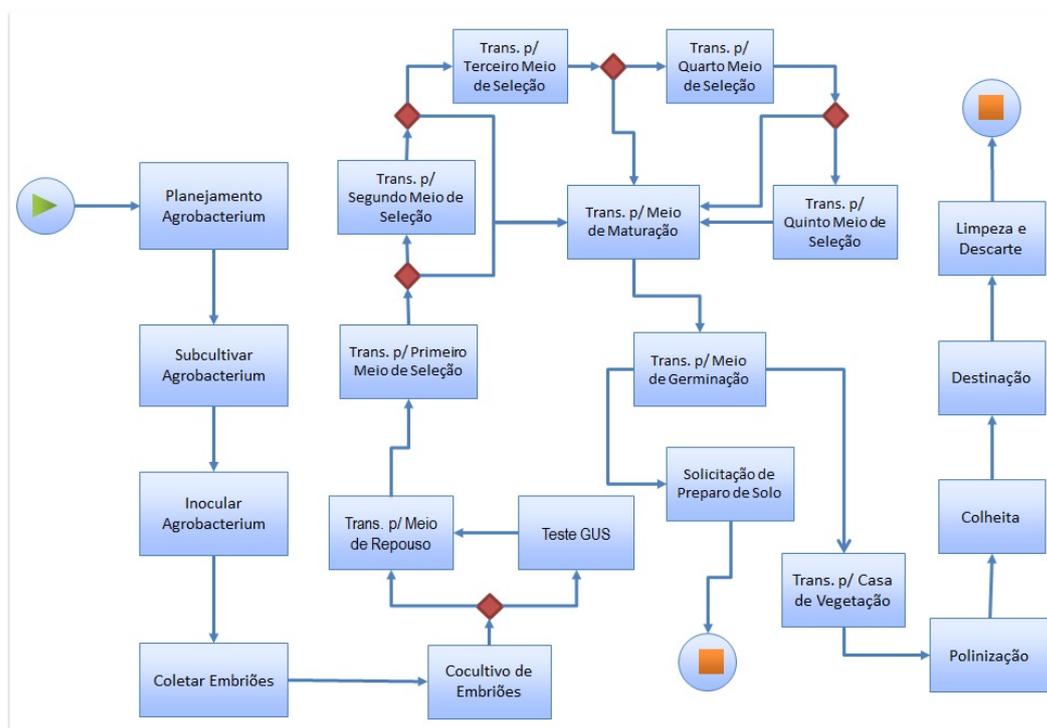


Figura 5.2. Fluxo de trabalho da transformação de milho e sorgo via *Agrobacterium tumefaciens*. Cada caixa representa uma atividade do processo e as setas de uma atividade para a outra representam uma transição.

A atividade inicial desse fluxo de trabalho é o Planejamento (Figura 5.3). Essa atividade é usada para registrar as informações gerais do experimento de transformação

5.1. FLUXO DE TRABALHO DA TRANSFORMAÇÃO VIA *Agrobacterium tumefaciens* 51

como o genótipo da planta, o número do experimento, o nome do usuário responsável, a construção gênica, etc. Um identificador do experimento é associado a cada experimento pelo sistema de forma que através desse identificador todas as informações relacionadas ao experimento sejam rastreáveis.

Figura 5.3. Tela do sistema *FluxTransgenics* demonstrando a primeira atividade do fluxo de trabalho da transformação via *Agrobacterium tumefaciens*. Nessa atividade, informações como a data, o responsável, o genótipo e a construção gênica são registrados.

Na atividade de planejamento, o registro da construção gênica que será usada é essencial. O sistema disponibiliza uma tela (Figura 5.4) que o usuário abre ao clicar em um botão e na tela o usuário é capaz de fazer uma busca por uma construção gênica previamente cadastrada no sistema. Após escolher a construção gênica, o sistema apresenta o volume disponível e assim o usuário entra com a quantidade de construção gênica que será usada no experimento. Caso o usuário solicite uma quantidade acima da disponível, o sistema alerta o mesmo sobre esse fato obrigando o usuário a mudar o valor. Em alguns casos, apenas um frasco de construção gênica não tem a quantidade de construção gênica o suficiente para realização do experimento, sendo assim, o usuário deve escolher mais de um frasco para obter a quantidade necessária. O sistema permite que sejam escolhidos vários registros para atender a esse cenário.

A atividade após o *Planejamento Agrobacterium* é a atividade *Subcultivar Agrobacterium* que por sua vez é seguida pela atividade *Inocular Agrobacterium*. Para essas atividades o LIMS grava a data em que a atividade está sendo executada e o protocolo onde estão os procedimentos para realização da atividade. Para todas as atividades anteriores e as seguintes, toda a informação relacionada a cada atividade é registrada no sistema como um conjunto de atributos. Na atividade *Coleta de Embriões*, a principal informação é a data de coleta dos embriões. Na atividade *Cocultivo de Embriões*, os

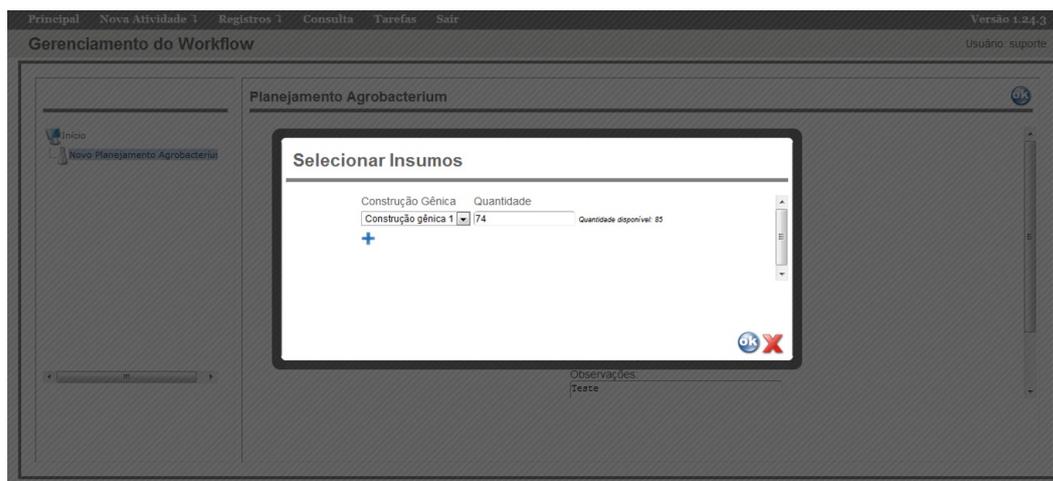


Figura 5.4. Tela do sistema *FluxTransgenics* que faz parte da atividade *Planejamento*. Esta tela é usada para escolher a construção gênica que será usada no experimento.

embriões coletados na atividade anterior são geneticamente modificados e as informações relacionadas são gravadas no sistema como o tamanho e quantidade dos embriões, o meio de cultura e o teste confirmatório (Figura 5.5). Assim como na atividade *Planejamento Agrobacterium* um insumo é escolhido (construção gênica), na atividade de cocultivo, um insumo também é associado à atividade (meio de cultura). O meio de cultura é uma solução preparada em laboratório e o registro desse preparo é feito na visão de controle de insumos através do fluxo de trabalho *Preparo de Solução*. A solução para representar o meio de cultura é escolhida em uma tela como a da (Figura 5.6). Assim como na escolha da construção gênica, o sistema valida a quantidade inserida e permite a escolha de múltiplos registros de solução.

Após a execução da atividade *Cocultivo de Embriões*, existem duas possibilidades de fluxo de trabalho. Em alguns casos a atividade *Teste GUS* é realizada para verificar a qualidade do meio de cultura e após esse teste a próxima atividade é executada. Em outros casos o *Teste GUS* não é realizado e a próxima atividade é a *Transferência para Meio de Repouso*. Nos experimentos de transformação de milho e sorgo, os embriões permanecem por um tempo em um meio de cultura e podem ser transferidos para cinco diferentes meios de seleção (meio de cultura). A transferência para cada meio constitui em uma nova atividade no fluxo com seus próprios atributos. Para cada transferência para um meio diferente, um controle dos calos gerados é realizado através do registro das entradas e saídas (visão de entradas e saídas). Essa informação é armazenada como número de identificação e número de cópias dos calos (Figura 5.9). Na transferência, os calos armazenados são identificados pelo número de identificação usado como entrada

5.1. FLUXO DE TRABALHO DA TRANSFORMAÇÃO VIA *Agrobacterium tumefaciens* 53

Árvore | Linear

Cocultivo de Embriões - 49

Transformação

Data:

Data da Captura da Bactéria:

Data da Estria da Colônia:

Absorbância Inicial:

Quantidade de Embriões:

Tamanho dos Embriões:

Absorbância no Momento da Transformação:

Pré Tratamento:

Alguns embriões > 1,5mm

Figura 5.5. Tela do sistema *FluxTransgenics* para a atividade *Cocultivo de Embriões*. Após essa atividade ser executada, informações sobre a absorbância, quantidade e tamanho dos embriões, teste confirmatório, protocolos, etc são armazenados.

Árvore | Linear

Cocultivo de Embriões - 49

Selecionar Insumos

| Insumo | Quantidade | Quantidade disponível |
|----------------|------------|-----------------------|
| Iodo Sanizante | 34 | 100.4500 |
| Iodo Sanizante | 16 | 95.4500 |

Adicionar POP

Adicionar POP

Foto:
Escolher arquivo | Nenhum arquivo selecionado

Figura 5.6. Tela do sistema *FluxTransgenics* para seleção de insumos na atividade *Cocultivo de Embriões*. Nesta atividade, após os embriões serem transformados, eles são colocados no meio de cultura feito a partir de uma solução preparada no laboratório. Essa tela é usada para escolher o insumo (solução) que constituirá o meio de cultura.

para a próxima atividade. Esse comportamento é repetido até a atividade *Transferência para Casa de Vegetação*. A atividade *Solicitação de Preparo de Solo* é paralela à *Transferência para a Casa de Vegetação*, nesta atividade o pesquisador registra uma solicitação de preparo de canteiros e/ou vasos para plantar os calos modificados. A atividade seguinte à *Transferência para a Casa de Vegetação* é a *Polinização* onde o tipo usado (autocruzamento, cruzamento e *sibling*) pode ser escolhido em um menu *drop-down*. As saídas dessa atividade são o identificador da espiga de milho ou sorgo e o número de sementes da mesma. Após a polinização a próxima atividade é a *colheita*. Para essa atividade a quantidade de sementes colhidas é armazenada como uma saída, assim como a data e o protocolo usado. A atividade *Destinação* permite que a data

e o protocolo de execução sejam armazenados, assim como a destinação das sementes (Exemplo: o local onde as mesmas são estocadas). Por fim, a atividade *Limpeza e Descarte* encerra o fluxo, nesta a data e o protocolo de execução são registrados e o usuário também é capaz de registrar quais plantas e sementes serão descartadas.

Nesse fluxo de trabalho, as atividades que compreendem desde a *Transformação para o Primeiro Meio de Seleção* até *Limpeza e Descarte*, são associadas à visão de entradas e saídas. Com as entradas e saídas o desenvolvimento da planta (milho ou sorgo) é acompanhado de forma que no final do fluxo de trabalho é possível identificar qual foi o embrião que gerou uma semente transgênica. Para esse fluxo, é possível escolher o tipo de entradas e saídas não-atômico (Figura 5.7), ou seja, parte de uma entrada pode ser usada por uma atividade e a outra parte por outra atividade diferente. Além disso, a saída da atividade é associada com a entrada da atividade anterior (Figuras 5.8 e 5.9). Há um menu *drop-down* com o nome das entradas e ao escolher uma entrada se faz o relacionamento dessa com a saída que está sendo cadastrada.

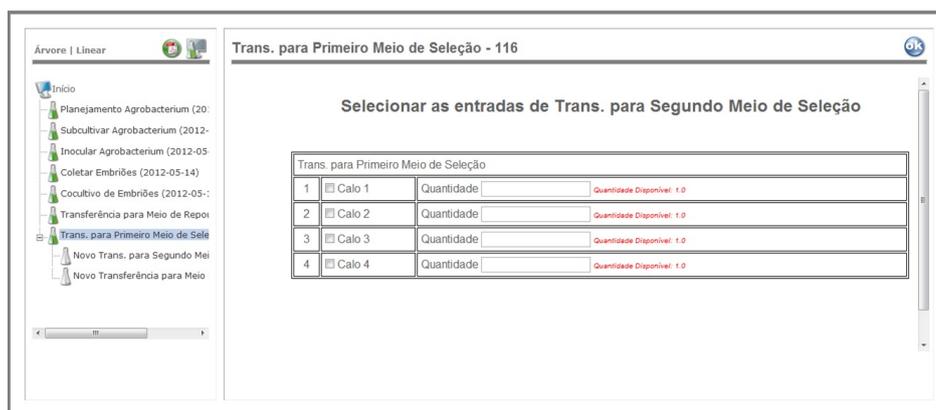


Figura 5.7. Tela do sistema *FluxTransgenics* mostrando entradas disponíveis para a atividade *Transferência para o Segundo Meio de Seleção*. Essas entradas são o resultado do cadastro das saídas na atividade anterior (*Transferência para o Primeiro Meio de Seleção*).

Para o sucesso do experimento de transformações via *A. tumefaciens* é fundamental que as atividades sejam executadas na data correta. Porém, são vários os experimentos que o pesquisador tem que lidar no dia-a-dia do laboratório e torna-se tarefa difícil lembrar quando cada atividade deve ser executada. Para resolver essa situação, a visão de agendamento permite que seja estipulada a agenda do fluxo de trabalho, dessa forma o sistema manterá uma lista com as atividades disponíveis para execução e, por conseguinte, o usuário será capaz de fazer pesquisas por data para identificar atividades disponíveis para execução. Com esse agendamento, quando o usuário executa atividades fora da data agendada (Figura 5.10), uma mensagem de alerta é apresentada para



Figura 5.8. Tela do sistema *FluxTransgenics* mostrando cadastro de uma saída da atividade *Transferência para o Segundo meio de Seleção* no qual é possível escolher a entrada relacionada.

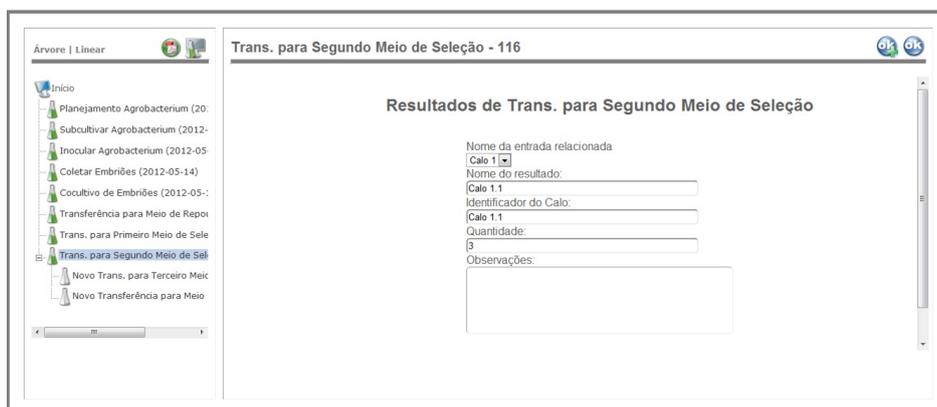


Figura 5.9. Tela do sistema *FluxTransgenics* de Cadastro de resultados (saída) da atividade *Transferência para o Segundo meio de Seleção*.

o mesmo, dessa forma, a responsabilidade de execução da atividade é do usuário e com o alerta o sistema evita que atividades sejam executadas fora da data prevista.

5.2 Fluxo de Trabalho da Transformação via Biobalística

O fluxo de trabalho da transformação via Biobalística (TVB) (Figura 5.11) tem modelagem similar a usada para o fluxo de trabalho da Transformação via *A. tumefaciens* (TAT) e compartilha algumas atividades deste como a Coleta de Embriões e a transferência para diferentes meios de seleção. A atividade *Planejamento Biobalística* desse fluxo é praticamente igual a atividade *Planejamento Agrobacterium* do fluxo TAT. A

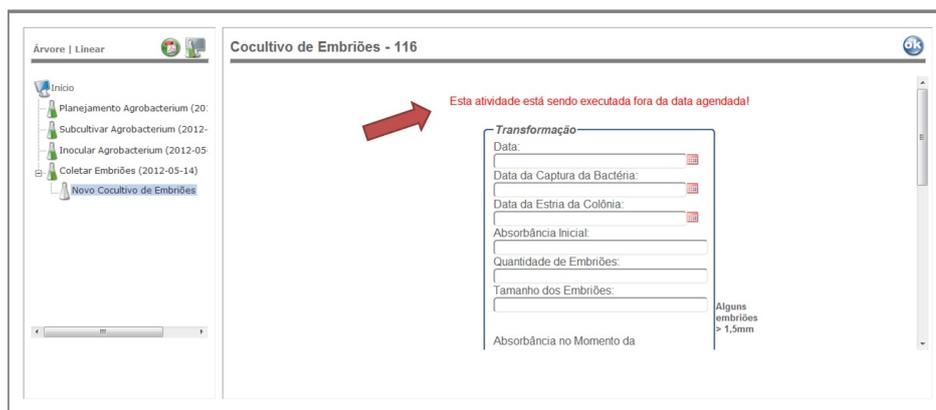


Figura 5.10. Tela do sistema *FluxTransgenics* mostrando mensagem apresentada quando o fluxo de trabalho tem a visão de *Agendamento* e a atividade é executada fora da data agendada.

principal diferença reside na construção gênica escolhida.

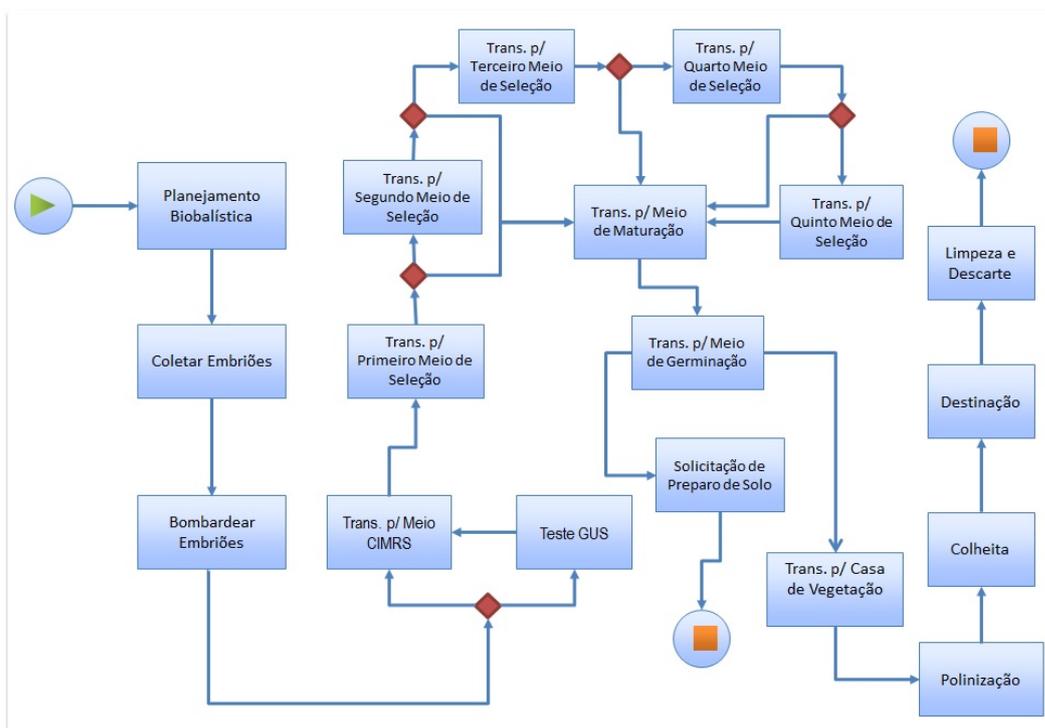


Figura 5.11. Fluxo de trabalho do processo de transformação de milho e sorgo via Biolística. Cada caixa representa uma atividade no processo e a seta de uma atividade para a outra representa uma transição.

Depois da transformação dos embriões, a principal diferença entre os dois fluxos é a transferência de explantes para o meio de cocultivo no caso da transformação via *A. tumefaciens* durante 3 dias ou a transferência dos embriões diretamente para o meio

de repouso no caso da transformação via Biobalística. A atividade *Bombardeamento de Embriões* é a principal atividade da transformação de milho e sorgo via biobalística (Figura 5.12). Para essa atividade, vários detalhes da transformação são gravados como a distância da membrana de retenção e explante, o número de tiros por placa, o volume da solução de micropartículas, a pressão do gás hélio, o tipo de partícula e informação sobre o teste confirmatório.

The screenshot displays the 'Gerenciamento do Workflow' interface. The top navigation bar includes 'Home', 'Nova Atividade', 'Registros', 'Consulta', 'Tarefas', and 'Logout', with 'Version 1.4' and 'Usuário: cristiano' on the right. The main content area is titled 'Bombardear Embriões - CS15092011-Bio'. On the left, a tree view shows 'Início', 'Planejamento (2011-09-15)', 'Coletar Embriões (2011-09-15)', and 'Novo Bombardear Embriões'. The main form, titled 'Transformação', contains the following fields:

| | |
|--|---------------------------|
| Data: | Ex: 01/01/2000 |
| Distância da Membrana de Retenção: | cm |
| Distância da Membrana de Explante: | cm |
| Número de Tiros Por Placa: | |
| Quantidade de Solução de Micropartículas: | mL/membrana carreadora |
| Data de Preparo da Solução de Micropartículas: | Ex: 01/01/2000 |
| Pressão do Gás Hélio: | PSI |

Figura 5.12. Tela do sistema *FluxTransgenics* com a atividade *Bombardear Embriões*. Através dessa atividade, é possível preencher todas as informações a cerca do bombardeamento de embriões, tal como o número de tiros por placa e a pressão do gás helium no momento do bombardeamento dos embriões.

Assim como o fluxo TAT, o fluxo TVB também tem as atividades a partir de *Trans. p/ Primeiro Meio de Seleção* associada à visão de entradas e saídas. Logo, esse fluxo também permite o rastreamento eficaz do desenvolvimento do embrião modificado geneticamente. Além disso, esse fluxo também tem o fluxo de trabalho *Agendamento* associado a ele, permitindo que ao iniciar um experimento de biobalística, uma agenda seja gerada e o usuário possa acompanhar as datas agendadas para execução na tela de tarefas do sistema.

Os fluxos construídos para representar os dois tipos de transformação abrangem o conjunto de atributos com o intuito de registrar e gerenciar todas as atividades realizadas durante o processo de transformação. Além disso, toda a informação armazenada pode ser facilmente recuperada através dos relatórios gerados pelo sistema.

5.3 Coleta de Embriões e Sub-cultivo de Calos

A cada três semanas, técnicos do laboratório da EMBRAPA Milho e Sorgo fazem a coleta de calos e embriões que serão usados na transformação via Biobalística. Uma vez coletado o embrião, o mesmo é colocado em meio de cultura e a cada três semanas é sub-cultivado ou transferido para outra placa até o limite de três meses. Durante o período de três meses, o calo poderá ser usado a qualquer momento em experimentos de transformação via Biobalística ou em testes moleculares. Caso os calos ou embriões não sejam usados, ao final de três meses são descartados.

Para modelar esse processo de forma a registrar todos os passos e dados inerentes à coleta de embriões e sub-cultivo de calos foi desenvolvido o fluxo de trabalho *Coleta de Embriões e Sub-cultivo de Calos* (Figura 5.13)

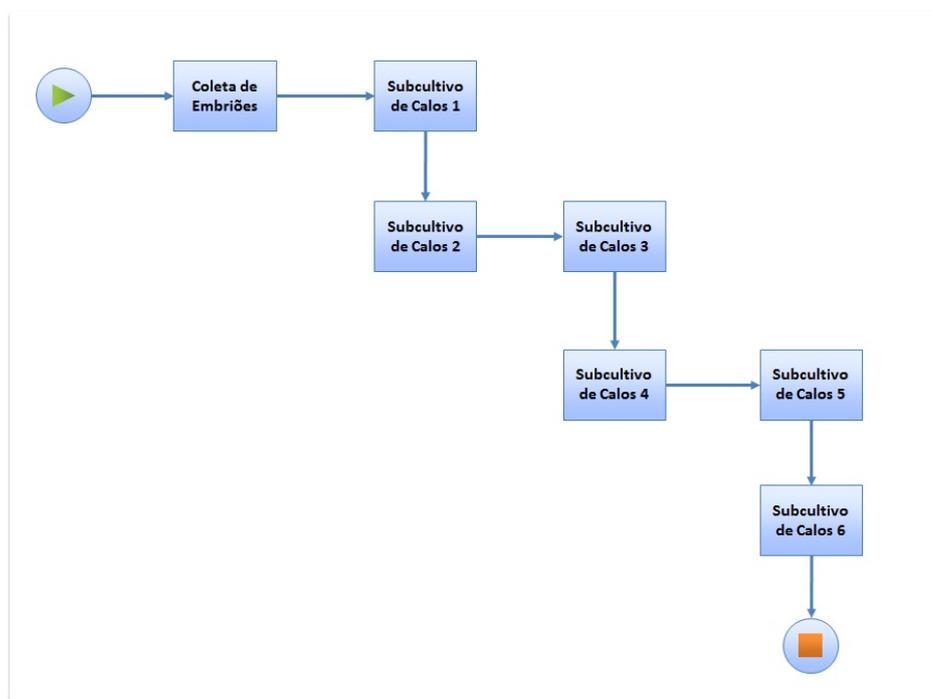


Figura 5.13. Modelagem do fluxo de trabalho *Coleta de Embriões e Sub-cultivo de Calos*.

A visão de agendamento está presente nesse fluxo de trabalho. Assim, ao executar o fluxo de trabalho *Agendamento*, quando se inicia a execução do fluxo *Coleta de Embriões e Sub-cultivo de Calos*, a agenda será gerada. Como mencionado anteriormente, esse fluxo acontece no intervalo de três meses, logo, o escalonamento das atividades acontece nesse intervalo. O agendamento é executado de forma que quando o usuário executar uma atividade fora da data agendada, um alerta é apresentado para o mesmo.

5.4 Sub-cultivo de Controle Positivo

A cada três semanas, assim como no processo de *Coleta de Embriões e Sub-cultivo de Calos*, é realizado o fluxo de trabalho *Sub-cultivo de Controle Positivo* (Figura 5.14). Esse é constituído de duas atividades: *Planejamento* e *Subcultivo de calos*. A primeira atividade tem o mesmo objetivo que a primeira do fluxo de transformação via *A. tumefaciens*, porém os atributos são diferentes, apesar dessa atividade também fazer uso de construção gênica. Os atributos são: fonte, data, número de placas, genótipo, além do protocolo. A segunda atividade tem uma característica peculiar, a seta presente nessa atividade dando a ideia de repetição indica que a atividade pode ser executada “infinitamente”. Essa característica existe em razão dessa atividade poder ser repetida a cada três semanas de acordo com a necessidade do usuário.

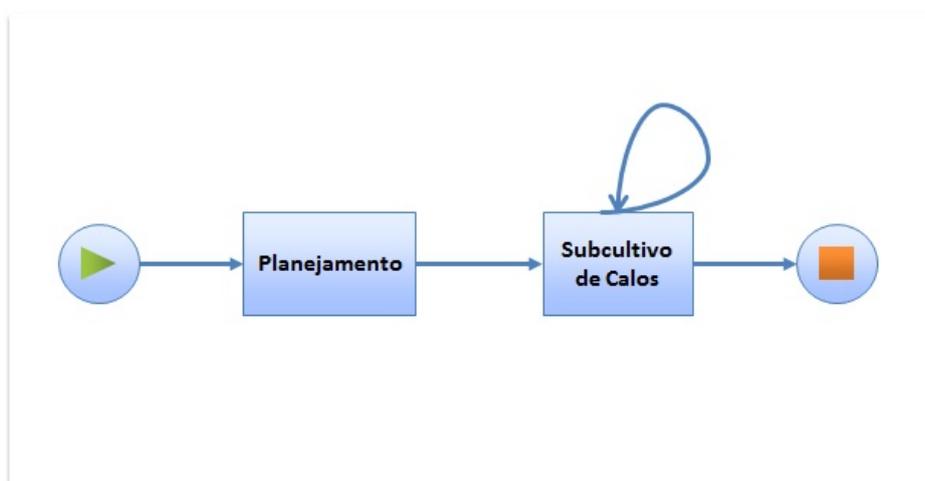


Figura 5.14. Modelagem do fluxo de trabalho *Sub-cultivo de Controle Positivo*.

5.5 Plantio de Sementes

O fluxo de trabalho *Plantio de Sementes* (Figura 5.15) tem como primeiro passo o registro da solicitação dos canteiros para plantio de sementes de milho ou sorgo. Esse registro é abstraído pela atividade *Solicitação de Preparo de Solo* e as informações requeridas por essa incluem a data e o responsável pela solicitação, o responsável pelo preparo, arquivo com as informações do balanço de nutrientes dentre outros. Uma semana após a solicitação do solo, é realizado o plantio e as informações são a data, o local de plantio, o protocolo e, nos casos necessários, o teste confirmatório. Após o plantio, as atividades são: polinização, colheita, destinação e limpeza e descarte. Essas atividades são similares as atividades dos fluxos de trabalho TAT e TVB.

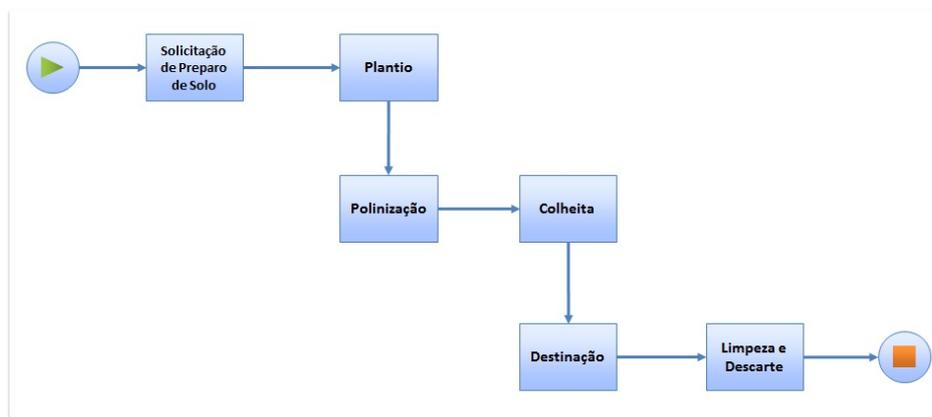


Figura 5.15. Modelagem do fluxo de trabalho *Plantio de Sementes*.

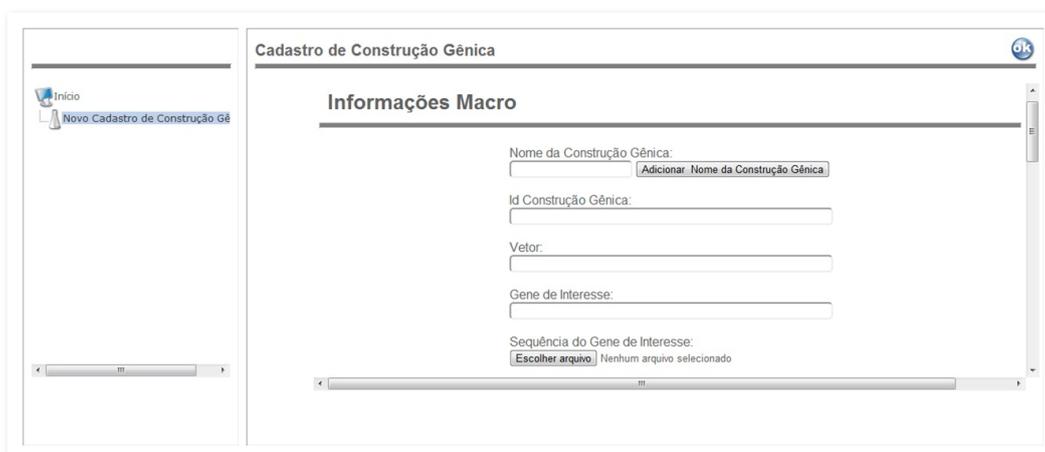
5.6 Controle de Insumos no *FluxTransgenics*

5.6.1 Cadastro de Construção Gênica

Um dos primeiros passos da transformação via *A. tumefaciens* é providenciar a construção gênica. Essa é um insumo que geralmente o laboratório da EMBRAPA adquire através da compra via fornecedores externos. Assim, é necessário que as construções adquiridas sejam devidamente cadastradas no sistema de forma que os pesquisadores tenham controle da quantidade e das construções existentes no laboratório. Sendo assim, desenvolveu-se o fluxo de trabalho cadastro de construção gênica. Esse cadastro faz parte da visão de controle de insumos e permite ao usuário o registro das informações inerentes à cada construção.

O sistema abstrai a construção gênica na perspectiva de dois elementos: informação macro e informação específica. A informação macro, vista no capítulo 4, se refere a um conjunto de objetos (frascos) físicos, ou seja, é a informação genérica desses objetos. No caso da construção gênica, a informação genérica (macro) (Figura 5.16) compreende dados como o nome da construção gênica, o gene de interesse, o nome da bactéria, arquivos (pdf, doc, txt, etc) com o diagrama do plasmídio, da construção gênica, foto do gel quando houver, protocolo da construção e a quantidade de itens da mesma. Como essa é a informação macro, a quantidade de itens se refere ao número de frascos da construção gênica. O cadastro da construção gênica se refere ao registro da informação de um frasco ou mais que existam fisicamente no laboratório. A informação específica (Figura 5.17) é referente aos frascos e os atributos desses são basicamente o nome da construção gênica, o status, o volume inicial e atual, se descartado ou não e a localização da construção gênica.

A informação macro tem um campo denominado *Quantidade de itens*, esse campo



The screenshot displays the 'Cadastro de Construção Gênica' interface. The main area is titled 'Informações Macro' and contains the following fields and controls:

- Nome da Construção Gênica:
- Id Construção Gênica:
- Vetor:
- Gene de Interesse:
- Sequência do Gene de Interesse: Nenhum arquivo selecionado

Figura 5.16. Tela do sistema *FluxTransgenics* para entrada da informação macro do cadastro de construção gênica.



The screenshot displays the 'Cadastro de Construção Gênica' interface. The main area is titled 'Informações Específicas' and contains the following fields and controls:

- Nome da Construção Gênica: Construção gênica 2
- Status: Estoque
- Volume Inicial: 200
- Volume Atual: 200
- Localização: EST-1-Pra/Gav/Por-8
- Descartado:

Figura 5.17. Tela do sistema *FluxTransgenics* para entrada das informações específicas que caracterizam cada frasco de construção gênica.

é preenchido com um valor numérico que indica a quantidade de frascos de construção gênica que está sendo cadastrado. Por exemplo, se são cinco frascos, então esse campo será preenchido com o valor cinco e dessa forma haverá cinco registros (itens) no banco de dados com a informação específica após a finalização do cadastro. Nesse exemplo, os cinco registros terão os campos com os mesmos valores, ou seja, o status, o volume inicial e final, e a localização de cada registro serão iguais e o que identificará cada registro será um identificador que o sistema atribuirá a cada um automaticamente. Mas, suponha que por algum motivo alguma informação dos cinco itens não seja igual para todos, por exemplo, a localização é diferente para cada item (registro). Então, é necessário discriminar essa informação de cada item, sendo assim o usuário pode abrir uma tela (Figura 5.18) onde serão listados os itens de acordo com a quantidade inserida

no campo *Quantidade de itens* da informação macro e assim os campos de cada registro (item) poderão ser alterados.

| | Nome da Construção Gênica: | Status: | Volume Inicial: | Volume Atual: | Localização: | Descartado: |
|----|----------------------------|---------|-----------------|---------------|---------------------|-------------|
| #1 | Construção gênica 2 | Estoque | 200 | 200 | EST-1-Pra/Gav/Por-8 | Não |
| #2 | Construção gênica 2 | Estoque | 200 | 200 | EST-1-Pra/Gav/Por-8 | Não |
| #3 | Construção gênica 2 | Estoque | 200 | 200 | EST-1-Pra/Gav/Por-8 | Não |
| #4 | Construção gênica 2 | Estoque | 200 | 200 | EST-1-Pra/Gav/Por-8 | Não |
| #5 | Construção gênica 2 | Estoque | 200 | 200 | EST-1-Pra/Gav/Por-8 | Não |

Figura 5.18. Tela do sistema *FluxTransgenics* que permite alterar os atributos específicos.

5.6.2 Cadastro de Reagente

Outro insumo utilizado nos experimentos dos laboratórios da EMBRAPA é o reagente. Os reagentes são as enzimas, soluções, suspensões, pós e similares utilizados no preparo de outras soluções utilizadas nas transformações, como por exemplo, meio de cultura para desenvolvimento dos embriões e calos modificados geneticamente. Os reagentes são em geral, insumos que o laboratório da EMBRAPA adquire através da compra via fornecedores externos.

Assim como a construção gênica, o reagente também é cadastrado no sistema de forma que o gerente do laboratório possa controlar o estoque desse insumo, podendo assim prever reposições de estoque. Assim como o cadastro de qualquer insumo no *FluxTransgenics*, o cadastro de reagentes também é feito na perspectiva das informações macro e específica. Os dados obrigatórios da informação macro do cadastro de reagentes são o nome do insumo, o lote do fornecedor, o nome da marca, a data de validade e a quantidade de itens (Figura 5.19). A informação específica de cada item por sua vez, se resume ao nome do insumo (reagente), o status, a localização, a medida, peso atual e inicial, volume inicial e atual e número de reações atual e inicial. Nem todos os reagentes tem a mesma forma física, ou seja, alguns reagentes podem ser na forma de pó, outros líquidos e outros podem ser referidos como reações. Dessa forma, quando um reagente for na forma de pó, sua medida será peso, quando líquido, volume e quando for reação, n° de reações. Para abstrair essa característica a visão de controle de

insumos permite que atributos de um formulário sejam disponíveis para preenchimento condicionados ao valor de outro campo. No caso do cadastro de reagente, os campos peso inicial e peso atual apenas serão visíveis quando o atributo medida for preenchido com peso (Figura 5.20). A mesma lógica se aplica para volume e reações. Assim, como no cadastro de construção gênica, no cadastro de reagentes também é possível discriminar a informação específica de cada item para aqueles atributos que diferem de um item para outro.



Figura 5.19. Tela do sistema *FluxTransgenics* para entrada da informação macro do cadastro de reagentes.

5.6.3 Cadastro de Solução

O cadastro de solução tem os mesmos elementos vistos nos cadastros de reagente e construção gênica. Ou seja, a informação macro e específica e a possibilidade de detalhar cada item nos casos em que há itens com atributos que diferem dos demais (informação específica). O fluxo cadastro de solução existe para registrar as soluções que são compradas ou doadas para o laboratório. Os principais atributos da informação macro são: nome do insumo, qualidade da água, ph, data de validade e protocolo. Os dados da informação específica são iguais aos do cadastro de reagente.

5.6.4 Cadastro de Receita

Uma das atividades dos laboratórios da EMBRAPA Milho e Sorgo é o preparo de soluções. Essas soluções são usadas para diversos fins, entre eles o preparo de meio de cultura onde os embriões geneticamente modificados ficam até se tornarem plantas e serem plantados na casa de vegetação. Para preparo de cada solução existe um

The figure consists of two screenshots, A and B, of a web application interface titled 'Cadastro de Reagente'. Both screenshots show a sidebar on the left with 'Inicio' and 'Novo Cadastro de Reagente' links. The main area is titled 'Informações Específicas' and contains several input fields: 'Nome Insumo:' (filled with 'Solução de HCL'), 'Id Insumo:' (filled with '9'), 'Status:' (dropdown menu with 'Estoque' selected), and 'Localização:'. In screenshot A, the 'Medida:' dropdown menu is open, showing options: 'Volume', 'Peso' (highlighted in blue), and 'Nº de Reações'. A red arrow points to the 'Medida:' dropdown. In screenshot B, the 'Medida:' dropdown is closed and 'Peso' is selected. Below it, the 'Peso Atual:' and 'Peso Inicial:' input fields are now visible and highlighted with a red rectangular box.

Figura 5.20. (A) Tela do sistema *FluxTransgenics* para detalhamento de informação específica. Na entrada das informações específicas, o campo medida é um *drop-down* menu com as opções peso, volume e nº de reações. A escolha de uma das opções resulta no aparecimento dos outros campos: peso inicial e atual se peso for escolhido; volume inicial e atual se volume for escolhido; e nº de reações inicial e atual caso o opção nº de reações seja selecionado. (B) Resultado da escolha do valor de medida peso. Ao escolher essa opção os campos Peso Inicial e Peso Atual foram habilitados para preenchimento.

documento com informações dos insumos que serão usados. Considerando que esse documento guia o pesquisador no momento da execução da atividade de preparo da solução e deve ser seguido em conformidade com as BPL foi necessária a construção de um fluxo de trabalho para Cadastro de Receita. Esse cadastro consiste do nome da receita, o *upload* do protocolo para preparo do insumo e a inserção dos nomes dos insumos necessários para o preparo (Figura 5.21).

O resultado da utilização da receita é a geração automática pelo sistema da tela de seleção dos insumos necessários para preparo da solução (Figura 5.22). É apresentada ao usuário a lista dos insumos necessários para o preparo (no exemplo da figura são *parafina pura*, *reativo mayer* e *guaiacol*), a quantidade do insumo que será usada e, para facilitar o trabalho do pesquisador quando o mesmo escolhe o insumo,

logo em seguida a quantidade disponível do insumo é apresentada, haja vista que não se deve inserir uma quantidade maior que a disponível. O botão “mais” da tela permite inserir mais entradas, possibilitando assim que seja selecionado mais de um insumo. O fato da receita apontar os insumos que devem ser escolhidos, não necessariamente implica que esses devam ser escolhidos, por exemplo, se a receita diz que o insumo *A* é necessária para o preparo, porém o laboratório não tem à disposição esse insumo mas tem o insumo *B* que é um substituto, então cabe ao usuário escolher o insumo *B* não causando prejuízo ao processo em questão. Ressalta-se que a receita é um guia para o usuário ajudando o mesmo a evitar possíveis erros de entradas de dados.

Figura 5.21. Tela do sistema *FluxTransgenics* para Cadastro de Receita. Na informação macro do cadastro se insere o nome da receita e se faz o upload de um arquivo com o protocolo. Na parte de informações específicas, permite-se o registro dos nomes dos insumos usados na receita.

A receita cadastrada no fluxo de trabalho Cadastro de Receita será usada no fluxo de trabalho de preparo de solução, e caso, posteriormente apareçam outros fluxos de trabalho para preparo de outros insumos, como por exemplo o insumo *I*, não é necessário criar outro fluxo de trabalho de Cadastro de Receita e sim, apenas cadastrar outra receita no qual o documento com os procedimentos para o preparo do insumo *I* será diferente assim como o nome da receita e os nomes dos insumos que serão usados no preparo desse insumo.

Um fator importante é a padronização do nome dos insumos o que impede o risco de se ter mais de um nome que se refere ao mesmo insumo, dessa forma optou-se por cadastrar esses nomes no banco de dados de forma que no cadastro de receita, reagente e solução e no preparo de solução, quando é necessário preencher algum campo que se refere ao nome de insumo, o sistema apresenta uma lista com esses nomes. Dessa

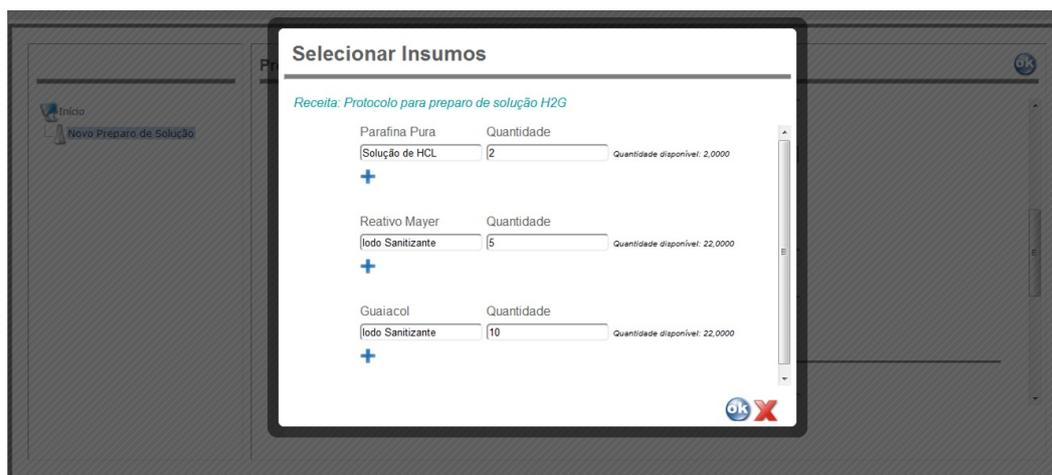


Figura 5.22. Tela do sistema *FluxTransgenics* demonstrando o uso da receita para preparo de solução. Os insumos relacionados à receita durante o cadastro da mesma são usados para gerar a tela onde o usuário seleciona os insumos usados no preparo.

forma, o usuário deve manter um catálogo com o nome dos insumos e atualizar quando necessário.

5.6.5 Preparo de Solução

Os experimentos de transformação requerem soluções que são usadas durante a fase de maturação dos embriões e calos de sorgo e milho geneticamente modificados. O preparo dessas soluções é crucial para a realização dos experimentos e manter informações sobre quem fez, onde foram feitas e como foram feitas as soluções atribuem maior confiança e qualidade às sementes transgênicas produzidas pelo laboratório. Logo, um fluxo de trabalho que permita o rastreamento da realização dessa atividade torna-se importante e pensando nisso se desenvolveu o fluxo de trabalho *Preparo de solução*. Esse fluxo é muito parecido com o fluxo *Cadastro de reagente* e *Construção gênica*, a diferença está no uso de uma receita onde o usuário escolhe aqueles insumos que juntos constituem a solução.

Assim como no *Cadastro de insumos*, no preparo se registra a informação genérica (informação macro) dos frascos (itens) que existem fisicamente no laboratório e a informação específica de cada um. A solução tem como principal informação macro o nome do insumo, data de preparo e validade da solução, volume de água utilizada, qualidade da água (que pode ser deionizada, destilada, bidestilada, tridestilada ou miliq) e o POP (Figura 5.23). O POP é protocolo de preparo (receita) e ao clicar no botão ao lado desse campo, o sistema abre uma tela como a da Figura 5.22, ou seja, de

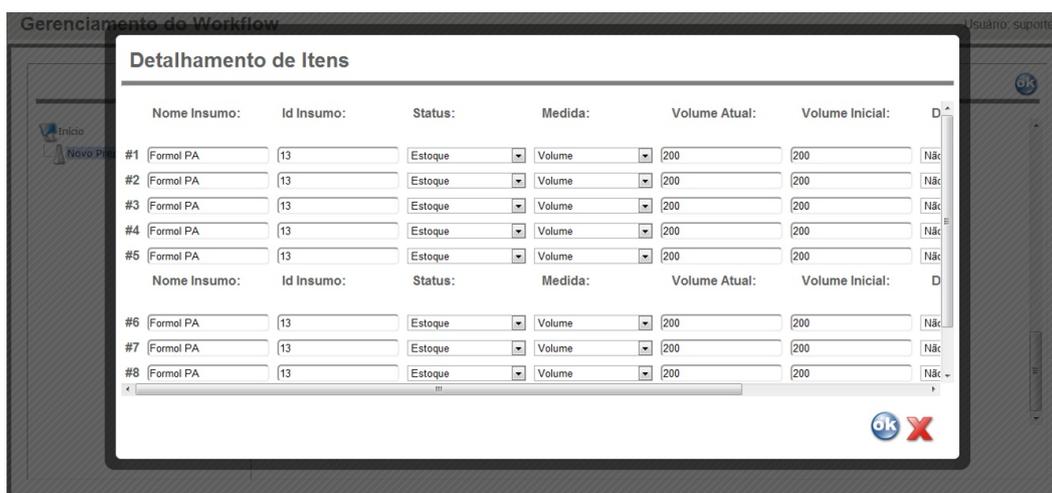
acordo com a receita escolhida para o campo POP, a tela onde se escolhe os insumos é renderizada conforme a lista dos insumos registrados no momento do cadastro da receita. A tela apresenta a lista de insumos que o usuário deve usar para preparar a solução.

Uma solução é preparada com reagentes e soluções, ou seja, usa-se uma solução preparada pelo fluxo *Preparo de solução* ou cadastrada pelo fluxo *Cadastro de solução* para preparar outra solução. As entradas da tela de seleção de insumos, onde o usuário escolhe o insumo que está cadastrado no banco de dados, fornece ao usuário todos os insumos armazenados no banco de dados, isto é, é apresentado para o usuário em uma mesma lista tanto os reagentes quanto as soluções.



The screenshot shows a web application window titled "Preparo de Solução". On the left, there is a sidebar with "Início" and "Novo Preparo de Solução" buttons. The main area contains a form with the following fields: "Data de Validade:" (text input), "PH:" (text input), "Volume da Água:" (text input), "Qualidade da Água:" (dropdown menu), "POP:" (text input with a "Selecionar Insumos" button), "Id POP:" (text input), and "Quantidade de Itens:" (text input). There is an "OK" button in the top right corner.

Figura 5.23. Tela do sistema *FluxTransgenics* com formulário para entrada de dados do preparo de solução.



The screenshot shows a web application window titled "Gerenciamento do Workflow" with a sub-window titled "Detalhamento de Itens". The table below lists items with columns for "Nome Insumo", "Id Insumo", "Status", "Medida", "Volume Atual", and "Volume Inicial". The table contains 8 rows of data, all showing "Formol PA" with "Id Insumo" 13 and "Volume Atual" 200. There is an "OK" button in the top right corner and a "Usuário: suporte" label in the top right corner.

| | Nome Insumo: | Id Insumo: | Status: | Medida: | Volume Atual: | Volume Inicial: | D |
|----|--------------|------------|---------|---------|---------------|-----------------|----|
| #1 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #2 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #3 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #4 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #5 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| | Nome Insumo: | Id Insumo: | Status: | Medida: | Volume Atual: | Volume Inicial: | D |
| #6 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #7 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |
| #8 | Formol PA | 13 | Estoque | Volume | 200 | 200 | Nã |

Figura 5.24. Tela do sistema *FluxTransgenics* do detalhamento de itens do fluxo de trabalho *Preparo de solução*.

A informação específica tem os mesmos atributos que a informação específica do cadastro de solução e reagente. Ou seja, o *status*, o nome do insumo, a medida, a localização e o comportamento onde de acordo com a medida escolhida é que os demais atributos aparecem. Também é possível discriminar os atributos da informação específica como no cadastro de reagentes com a tela de detalhamento de itens (Figura 5.24)

5.6.6 Visualização de Reagentes, Solução e Construção Gênica

Uma das questões relevantes na gestão de laboratórios é o controle dos insumos utilizados nos experimentos. Assim, é de grande interesse dos gestores dos laboratórios ter informações precisas sobre a situação dos insumos mantidos nas dependências do laboratório. Estas informações são desde dados como a quantidade disponível, o histórico de utilização dos insumos, quando será necessário repor algum até o número de experimentos que é possível executar com um dado insumo. Possuir estas informações permite ao gerente organizar melhor o laboratório, inclusive, evitando desperdícios e ineficiência. Por exemplo, suponha o cenário em que insumos próximos da data de validade não são utilizados por desconhecimento que o insumo está prestes a se tornar impróprio para uso. Isso obriga o descarte do insumo e conseqüentemente acarreta no desperdício de recursos financeiros. Para permitir o acompanhamento dos insumos registrados no sistema, foi desenvolvido um fluxo de trabalho para visualização dos insumos usados em cada fluxo de trabalho dos laboratórios da EMBRAPA: construção gênica, reagente e solução.

Filtrar

1

Id (Inicial) Id (Final) Nome Insumo Id Insumo (Inicial) Id Insumo (Final)

Data (Inicial) Data (Final) Lote do Fornecedor Marca Data de Validade (Inicial)

Data de Validade (Final) Quantidade de Itens (Inicial) Quantidade de Itens (Final)



Registro(s)

2 3

| Marcar Todos <input type="checkbox"/> | Id | Nome Insumo | Id Insumo | Data | Lote do Fornecedor | Marca | Data de Validade | Quantidade de Itens | Total Estoque | Lista de Itens | Editar/Apagar |
|---------------------------------------|----|------------------|-----------|------------|--------------------|-------|------------------|---------------------|---------------|---|---|
| <input type="checkbox"/> | 1 | Iodo Sanitizante | 11 | 22/04/2012 | 0112 | UFMG | 22/04/2012 | 20 | 1782,6800 |  |   |
| <input type="checkbox"/> | 2 | Solução de HCL | 9 | 13/05/2012 | | | 28/05/2012 | 2 | 4,0000 |  |   |
| <input type="checkbox"/> | 3 | Orto Tolidina | 14 | 19/05/2012 | 0001 | UFMG | 15/06/2012 | 5 | 0 |  |   |

Figura 5.25. Tela do sistema *FluxTransgenics* para pesquisa de reagentes. 1: Filtro que permite a pesquisa dos insumos de acordo com um critério estabelecido. 2: Botão da interface para emissão de relatórios em formato pdf. 3: Botão da interface para emissão de relatórios em formato csv.

Registro(s)

| Marcar Todos | Id | Nome Insumo | Id Insumo | Data | Lote do Fornecedor | Marca | Data de Validade | Quantidade de Itens | Total Estoque | Lista de Itens | Editar/Apagar |
|--------------------------|----|----------------------|-----------|------------|--------------------|-------|------------------|---------------------|---------------|----------------|---------------|
| <input type="checkbox"/> | 1 | Iodo Sanitizante | 11 | 22/04/2012 | 0112 | UFMG | 22/04/2012 | 20 | 1782,6800 | | |
| <input type="checkbox"/> | 2 | Solução de HCL | 9 | 13/05/2012 | | | 28/05/2012 | 2 | 4,0000 | | |
| <input type="checkbox"/> | 3 | Orto Tolidina | 14 | 19/05/2012 | 0001 | UFMG | 15/06/2012 | 5 | 0 | | |
| <input type="checkbox"/> | 4 | Tintura de Iodo | 23 | 19/05/2012 | 0001 | UFMG | 05/10/2012 | 4 | 0 | | |
| <input type="checkbox"/> | 5 | Percloroato de Ferro | 25 | 19/05/2012 | 0004 | UFMG | 19/05/2012 | 3 | 0 | | |
| <input type="checkbox"/> | 6 | Iodo Sanitizante | 11 | 19/05/2012 | 0001 | UFMG | 31/05/2012 | 3 | 66,0000 | | |
| Total: | | | | | | | | 37 | 1852,68 | | |

Figura 5.26. Tela do sistema *FluxTransgenics* com lista dos reagentes. A visualização fornece o total de cada reagente e o total global em estoque que compreende todos os reagentes.

Nos fluxos *Visualização de Construção Gênica*, *Reagente* e *Solução*, todos os insumos registrados são listados. No caso do fluxo de *Visualização de Construção Gênica*, todas as construções gênicas são listadas, na *Visualização de Reagentes* (Figuras 5.25 e 5.26), todos os reagentes são listados e na *Visualização de solução*, é possível visualizar todas as soluções. Esses fluxos de trabalho permitem ajustes no estoque de insumos, ou seja, é possível editar um insumo. Para fazer isso o usuário deve ter perfil de administrador, caso contrário essa funcionalidade não é disponibilizada. Isso acontece porque a alteração do estoque é algo crítico, então somente é permitido tal ação por usuários com privilégios específicos.

5.7 Conclusão

O LIMS *FluxTransgenics* é a otimização do *framework* Flux para atender o laboratório da Embrapa Milho e Sorgo. Isso é um exemplo prático da flexibilidade do Flux baseado em fluxo de trabalho. Além disso, o *FluxTransgenics* mostra a aplicabilidade das visões com sucesso uma vez que foi possível a modelagem dos processos de produção de sementes transgênicas considerando o **Agendamento**, as **Entradas e Saídas** e o **Controle de Insumos**.

Capítulo 6

Conclusão

A quantidade de dados gerados por processos experimentais apresenta-se como desafio frente aos responsáveis por gerenciá-los. O desafio está naquilo que se refere ao armazenamento e disponibilização desses dados. Além disso, não só a quantidade em larga escala dos dados, mas a quantidade de processos com que um pesquisador lida no dia-a-dia de um laboratório torna desafiador garantir a qualidade dos dados e processos que fazem parte de um laboratório. A fim de garantir que os resultados gerados em laboratório são de qualidade e passível de auditorias, é comum que laboratórios de diversas áreas de atuação adquiram Sistemas de Gerenciamento de Informações de Laboratório (LIMS). Esses sistemas são softwares complexos com a capacidade de armazenamento e gerenciamento das informações produzidas em laboratório. Além disso, permite auditorias nos procedimentos executados, possui sistema de *backups*, capacidade de integrar com sistemas externos e interface amigável com o usuário. Porém, para um laboratório adquirir um LIMS, deve dispor de altos recursos financeiros. E um mesmo LIMS não é capaz de atender laboratórios de áreas diferentes, portanto, o Laboratório de Universalização do Acesso (LUAR) desenvolveu o LIMS SIGLa e em seguida desenvolveu o LIMS Flux tendo o código fonte do SIGLa como ponto de partida. O Flux passou, então, a adquirir funcionalidades não presentes no seu precursor (SIGLa), principalmente as visões.

O objetivo do trabalho em questão foi o desenvolvimento de um novo componente no *framework* Flux: visões. Isto é, formas diversas de compreender os dados de um fluxo de trabalho experimental que não seguem o fluxo principal e também podem ser agrupamentos de alguns dados experimentais (subconjunto desses dados) que tem alguma propriedade em comum que não necessariamente está diretamente relacionada ao fluxo de trabalho. E paralelamente ao desenvolvimento das visões, a modelagem do *Laboratório de Biologia Celular e Cultura de Tecidos* da Embrapa Milho e Sorgo (MG

- Brasil) para construção dos fluxos de trabalho dos processos desse laboratório de tal forma a otimizar o Flux e gerar o LIMS *FluxTransgenics*. Os objetivos citados foram alcançados uma vez que se desenvolveu as visões **Controle de Insumos, Entradas e Saídas** e **Agendamento**. Além do mais, a fim de verificar a aplicabilidade dessas visões, desenvolveu-se o LIMS *FluxTransgenics* considerando o desenvolvimento de fluxos de trabalhos que fazem uso das visões. O *FluxTransgenics* gerencia os processos do laboratório da Embrapa Milho e Sorgo para produção de organismos geneticamente modificados (organismos transgênicos). Os processos gerenciados pelo *FluxTransgenics* são: transformação via *Agrobacterium tumefaciens*, transformação via Biobalística, Coleta e Sub-cultivo de Calos, Sub-cultivo de Controle Positivo, Plantio de Sementes e Controle de Insumos.

6.1 Validação em Laboratório Real

O trabalho desenvolvido foi realizado em cooperação com os pesquisadores do *Laboratório de Biologia Celular e Cultura de Tecidos* da Embrapa Milho e Sorgo (MG - Brasil) de tal forma que esses pesquisadores colaboraram com a validação dos fluxos de trabalho desenvolvidos. Em um primeiro momento, uma versão do *FluxTransgenics* com os fluxos de trabalho das transformações via *Agrobacterium tumefaciens* e Biobalística foi desenvolvida e instalada para uso do laboratório. Algumas não-conformidades foram notadas pelos pesquisadores sendo que estas contribuíram para a versão final do sistema relatado nesse trabalho. A versão final está instalada e em uso pelos colaboradores da Embrapa Milho e Sorgo.

6.2 Trabalhos Futuros

As visões projetadas e implementadas, como relatado, nesse trabalho tornam o *framework* Flux robusto e capaz de cobrir importantes aspectos do gerenciamento das informações de laboratório. Porém, o Flux ainda não é capaz de receber uma modelagem de um fluxo de trabalho em formato *XPDL* que contemple mineração automática de dados. Por exemplo, o caso de estudo apresentado no trabalho mostra os fluxos de trabalho para gerenciar os experimentos de transformação para produção de sementes de milho e sorgo geneticamente modificadas. A expectativa é que em um período de anos de uso do sistema haverão muitos dados relacionados com os processos geridos. Então, é fundamental que o LIMS seja capaz de buscar padrões nesses dados de forma

a sugerir melhorias nos processos do laboratório. Logo, capacitar o Flux para lidar com mineração de dados é uma contribuição para o trabalho desenvolvido.

Para demonstrar como é possível a aplicação das visões em um cenário real, as mesmas foram empregadas para o desenvolvimento do *FluxTransgenics*. Mas, até o momento esse LIMS é o único exemplo prático do uso das visões. Como trabalho futuro se pode otimizar o Flux fazendo uso das visões. Essa otimização seria voltada para o gerenciamento da produção de outros organismos geneticamente modificados como o arroz e o feijão. Além disso, trabalhos podem ser desenvolvidos para atender outras áreas diversas uma vez que o presente trabalho teve como casos de estudo os processos de produção de transgênicos.

Referências Bibliográficas

- Abhilash, M. (2009). Applications of proteomics.
- Balaji, J.; Reddy, P.; Leeladevi, Y.; Crouch, J.; Mahalakshmi, V.; Buhariwalla, H.; Eshwar, K.; Mace, E.; Folksterma, R.; Senthilvel, S.; Varshney, R.; Kannan, S.; Rajalakshmi, R.; Prasanth, V.; Chandra, S.; Swarupa, L.; SriKalyani, P. & Hoisington, D. (2006). Laboratory information management software for genotyping workflows: applications in high throughput crop genotyping. *BMC Bioinformatics*, 7:383--388.
- Bensmail, H. & Haoudi, A. (2003). Postgenomics: Proteomics and bioinformatics in cancer research. *Journal of Biomedicine and Biotechnology*, p. 217.
- Coalition, T. W. M. (2012). <http://xml.coverpages.org/xpdl20010522.pdf>. Acessado em 01 de Janeiro de 2012.
- Computing Solutions, I. (2010). <http://www.labsoftlims.com>. Acessado em 03 de Outubro de 2010.
- Draghici, S. (2003). *Data Analysis Tools for DNA Microarrays*. Chapman & Hall/CRC.
- Editor, T. X. W. (2010). <http://www.together.at/prod/workflow/twe>. Acessado em 01 de Outubro de 2010.
- Elmasri, R. & Navathe, S. (2011). *Sistemas de banco de dados*. São Paulo: Pearson Addison-Wesley.
- Embrapa (2010). <http://www.cnpmc.embrapa.br/>. Acessado em 01 de Outubro de 2010.
- Embrapa (2012). <http://www.ctaa.embrapa.br/projetos/bpl/apresentacao.php>. Acessado em 01 de Janeiro de 2012.
- Estevam dos Santos, P. (2010). Boas práticas de laboratório (bpl) uma questão de qualidade. *Revista Intertox de Toxicologia, Risco Ambiental e Sociedade*, 3.

- Foundation, T. A. S. (2012). <http://tomcat.apache.org/>. Acessado em 04 de Novembro de 2012.
- Fromm, M. E.; Morrish, F.; Armstrong, C.; Williams, R.; Thomas, J. & Klein, T. M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Nature Biotechnology*, 8:833--839.
- Gelvin, S. B. (2003). Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews*, 67(1):16--37.
- Gimeno, C. (2003). Sistemas de gestión de la calidad en los laboratorios clínicos: certificación y acreditación. *Enfermedades Infecciosas y Microbiología Clínica*, 21(2):17-23.
- Gordon-Kamm, W. J. (1990). Transformation of maize cells and regeneration of fertile transgenic plants. *The Plant Cell*, 12(7):603--618.
- Hinton, M. D. (1995a). *Laboratory Information Management Systems: development and implementation for a quality assurance laboratory*. MARCEL DEKKER, New York, New York.
- Hinton, M. D. (1995b). *Laboratory Information Management Systems: development and implementation for a quality assurance laboratory*. MARCEL DEKKER, New York, New York.
- Inc., T. S. (2011). <http://www.trademarkia.com/villager-transgenic-animal-management-system-78363691.html>. Acessado em 01 de Setembro de 2012.
- Inmetro (2012). http://www.inmetro.gov.br/sidoq/arquivos/dicla/nit/nitdicla-35_02.pdf. Acessado em 01 de Janeiro de 2012.
- Klein, T. M.; Wolf, E. D.; Wu, R. & Sanford, J. C. (1987). High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature*, 327:70--73.
- Kohl, K. & Gremmels, J. (2010). Documentation system for plant transformation service and research. *Plant Methods*, 6(1):4.
- LabVantage Solutions, I. (2010). <http://www.sqllims.com>. Acessado em 03 de Outubro de 2010.
- MySQL (2012). <http://www.mysql.com/>. Acessado em 04 de Novembro de 2012.

- Nigel, J. T. & Claude, M. F. (2002). Microparticle bombardment as a tool in plant science and agricultural biotechnology. *DNA Cell Biol*, 21:963--977.
- Olivares, I. R. B. (2006). *Gestão de Qualidade em Laboratórios*. Editora Átomo, Campinas, São Paulo.
- Oracle (2012). <http://www.oracle.com/technetwork/java/javase/jsp/index.html>. Acessado em 04 de Novembro de 2012.
- Panisko, E. A., C. T. P. G. M. & Veenstra, T. D. (2002). The postgenomic age: Characterization of proteomes. *Exp. Hematol*, p. 97.
- Rasmussen, L.; Maddox, C. B.; Harten, B. & White, E. L. (2007). A successful lims implementation: Case study at southern research institute. *Journal of the Association for Laboratory Automation*, 12:384--390.
- Sanford, J.; Smith, F. & Russell, J. (1993). Optimizing the biolistic process for different biological applications. *Academic Press*, 217:483--509.
- Sanford, J. C. (1988). The biolistic process. *Trends in Biotechnology*, 6(12):299--302.
- Simões, A.; Faria Campos, A. C.; Delaat, D. M.; Abreu, V. & Campos, S. V. A. (2010). Sigla: An adaptable lims for multiple laboratories.
- Solutions, L. L. (2010). <http://www.labware.com/lwweb.nsf>. Acessado em 03 de Outubro de 2010.
- Teamsolutions, T. (2012). <http://www.together.at/prod/workflow/twe/manual>. Acessado em 01 de Janeiro de 2012.
- WfMC (2010). <http://www.wfmc.org/>. Acessado em 01 de Outubro de 2010.
- Yuji, I.; Hideaki, S.; Shozo, O.; Yukoh, H.; Toshihiko, K. & Takashi, K. (1996). High efficiency transformation of maize (*zea mays* l.) mediated by agrobacterium tumefaciens. *Nature Biotechnology*, 14:745--750.
- Yukoh, H.; Shozo, O.; Toshihiko, K. & Takashi, K. (1994). Efficient transformation of rice (*oryza sativa* l.) mediated by agrobacterium and sequence analysis of the boundaries of the t-dna. *The Plant Journal*, 6:271--282.

